

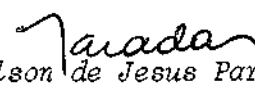


1. Publicação nº <i>INPE-2731-PRE/313</i>	2. Versão	3. Data <i>Maio, 1983</i>	5. Distribuição <input type="checkbox"/> Interna <input checked="" type="checkbox"/> Externa <input type="checkbox"/> Restrita
4. Origem <i>DIN/DPI</i>	Programa <i>IMAGE/INFORMÁTICA</i>		
6. Palavras chaves - selecionadas pelo(s) autor(es) <i>PROCESSAMENTO DE IMAGENS RECONHECIMENTO DE PADRÕES VISÃO POR COMPUTADOR</i>			
7. C.D.U.: <i>621.376.5:681.3.01:519.2</i>			
8. Título <i>PROCESSAMENTO DE IMAGENS, CLASSIFICAÇÃO DE PADRÕES E VISÃO ROBÓTICA</i>		10. Páginas: <i>15</i>	
		11. Última página: <i>13</i>	
9. Autoria <i>Nelson D.A. Mascarenhas Flávio R.D. Velasco</i>		12. Revisada por  <i>Ubirajara M.B. Lima</i>	
Assinatura responsável 		13. Autorizada por  <i>Nelson de Jesus Parada Diretor</i>	
14. Resumo/Notas <i>Este trabalho é um sumário dos principais aspectos de atividades correntes nas áreas de Processamento de Imagens, Classificação de Padrões e Visão Robótica.</i>			
15. Observações <i>Trabalho submetido para apresentação no 1º Congresso Nacional de Indústria - CONAI, no Maksoud Plaza Hotel-SP, de 11 a 15/07/83.</i>			

ABSTRACT

This article is a summary of the main aspects of current activities in the areas of Image Processing, Pattern classification and Robotics vision.

PROCESSAMENTO DE IMAGENS, CLASSIFICAÇÃO DE PADRÕES E VISÃO ROBÓTICA

Nelson D.A. Mascarenhas, Flávio R.D. Velasco

Instituto de Pesquisas Espaciais - INPE
Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
Caixa Postal 515 - São José dos Campos

RESUMO: Este trabalho é um sumário dos principais aspectos de atividades correntes nas áreas de Processamento de Imagens*, Classificação de Padrões* e Visão Robótica.

1 - PROCESSAMENTO DE IMAGENS

1.1 - INTRODUÇÃO

A área de Processamento de Imagens tem apresentado grande desenvolvimento nos últimos 15 anos, como resultado direto da evolução da tecnologia digital. As aplicações abrangem uma ampla gama de domínios, incluindo, por exemplo, análise de imagens aeroespaciais para detecção de recursos naturais em sensoriamento remoto de previsão meteorológica; transmissão de imagens de televisão de fac-símile por via digital; análise de imagens biomédicas como radiografia ou tomografia computadorizada; automação industrial envolvendo o uso de sensores visuais em robôs para inspeção ou manufatura.

As técnicas de Processamento de Imagens resultam de várias disciplinas como Ótica, Neurofisiologia, Colorimetria, Engenharia Elétrica e Ciência de Computação. Este artigo tutorial abordará apenas os aspectos ligados das duas últimas áreas, e apresentará uma descrição sumária das técnicas computacionais, sem maior detalhamento matemático. O leitor poderá encontrar maiores detalhes na bibliografia citada no final deste artigo e que serviu de base para a elaboração das suas duas primeiras partes. A exposição será orientada para os aspectos algorítmicos e não abordará os sistemas de "hardware" como sensores, arquitetura dos computadores ou visualizadores. Além disso é conveniente frisar que certas técnicas podem ser implementadas por processamento ótico, que não será coberto por este tutorial.

1.2 - AMOSTRAGEM E QUANTIZAÇÃO

Para processar uma imagem por computador, a função de duas variáveis que caracteriza a luminância é normalmente amostrada numa gra-

de regular, e os valores resultantes são discretizados num número finito de níveis de cinza. Imagens a cores requerem três componentes (por exemplo: azul, verde e vermelho), que são amostradas e quantizadas de maneira análoga. Em muitos casos, uma imagem consiste numa matriz de 512 x 512 elementos denominados "pixels" (abreviatura do inglês "picture elements"), cada um possuindo 64 níveis de cinza. Também é usual trabalhar com imagens binárias, como em aplicações industriais ou em fac-símile de documentos.

Se a imagem original apresenta detalhes muito finos, esta informação de alta frequência espacial será perdida se a amostragem tiver grande espaçamento. De acordo com o teorema da amostragem, a mínima taxa de amostragem necessária para a reprodução fiel de um sinal é o dobro da máxima frequência nele contida. A amostragem insuficiente de sinais pictóricos causa o chamado fenômeno de "aliasing", que provoca o aparecimento de estruturas espúrias na imagem, denominadas padrões de Moiré.

A quantização dos elementos resultantes da amostragem pode ser feita de maneira não uniforme ("tapered quantization"). Pelo menor espaçamento dos níveis de cinza na faixa em que estes ocorrem mais frequentemente, é possível minimizar o erro de quantização. Um número insuficiente de níveis de cinza irá provocar o aparecimento dos chamados falsos contornos, que apresentam transições bruscas de tonalidade, que originalmente eram suaves.

1.3 - REALÇAMENTO

As técnicas de realçamento de imagens são utilizadas, em geral, para melhorar a qualidade da imagem sob critérios subjetivos de avaliação do observador humano. Podem também servir como uma etapa de pré-processamento

* Uma versão preliminar do tutorial sobre as duas primeiras áreas foi apresentada no II Simpósio Brasileiro de Sensoriamento Remoto realizado em Brasília, em maio de 1982.

para sistemas de reconhecimento de padrões.

É frequente ocorrer a situação em que, após a digitalização da imagem (amostragem e quantização), os níveis de cinza ocupem apenas uma pequena parte da variação disponível. Por um processo de modificação da escala de cinza, é possível espalhar a distribuição desses níveis, tornando a informação existente mais visível ao olho humano. Tais métodos são também chamados de manipulação do contraste e envolvem o mapeamento linear ou não-linear (por exemplo: utilizando saturação) dos níveis originais, através do uso de tabelas.

A redistribuição dos níveis de cinza pode ter também como objetivo uma alteração do histograma da imagem, de modo a que este adquira uma forma desejada, que poderá ser umiforme ou ainda exponencial ou hiperbólica, levando em conta as características do sistema visual humano.

Uma imagem é frequentemente afetada por ruído ou interferência proveniente de várias fontes. Uma possível abordagem deste problema envolve a aplicação de técnicas estatísticas de estimação e filtragem ótima. Tais métodos exigem, em geral, uma caracterização da imagem original e da fonte de ruído por descrições probabilísticas, através de campos aleatórios, levando ao projeto dos filtros de Wiener ou de Kalman (filtragem recursiva). Entretanto, tais caracterizações são, por vezes, difíceis de serem obtidas e na prática, recorre-se então a técnicas de filtragem *ad hoc*. Assim, processos de limpeza de ruído podem ser aplicados envolvendo a imagem original com máscaras de coeficientes que possuam a característica de atenuar as altas frequências, onde o ruído é geralmente preponderante sobre o sinal, ou então no caso mais simples substituindo o ponto central de uma vizinhança 3×3 ou 5×5 "pixels" de uma imagem pela média nessa vizinhança. O preço que se paga nesse processamento é, em geral, uma suavização das variações da imagem. É possível aliviar este problema por um tipo de filtragem não-linear denominado filtro de mediana em que, numa janela com número ímpar de "pixels", o ponto central é substituído pela mediana dos valores observados.

Técnicas de realce de bordas tendem a aumentar a visibilidade de contornos de baixo contraste na imagem e a percepção de detalhe. Uma das operações mais simples para conseguir tal efeito é subtrair da imagem uma versão borrada dela mesma. A idéia subjacente é que essa versão contém essencialmente as componentes de baixa frequência da imagem, do que resulta uma filtragem passa-altas. Tal método é conhecido como "unsharp masking" e é utilizado por fotografos em seus processos de revelação química. Outras possíveis operações de realçamento de bordas consistem na aplicação de filtragem passa-altas por meio de máscaras convolucionais. É possível obter aproximações digitais dos op

radores gradiente ou laplaciano que são utilizadas também em detecção de bordas, a ser tratada com mais detalhes na secção sobre análise de imagens.

Através de um modelo em que a luminância da imagem é encarada como o produto de um fator de iluminação por outro de reflectância, pode-se realçar este último pela utilização das técnicas de filtragem homomórfica, que podem incorporar também modelos estatísticos para os processos envolvidos. Nesse ponto, a área de processamento de imagens tangencia a de filtragem digital bidimensional, que tem sido objeto de intensas investigações nos últimos anos.

A habilidade do olho humano em distinguir um número muito maior de cores do que tons de cinza conduziu ao desenvolvimento de técnicas de realçamento a pseudo-cor, onde os tons de uma imagem em branco e preto são mapeados para o espaço de cores.

Muitos sistemas óticos exibem pronunciadas distorções geométricas. Tal é o caso de câmeras de "vidicon" ou rastreadores multiespectrais a bordo de satélites, cuja atitude em relação à Terra é variável. Surge, portanto a necessidade de se corrigir tais distorções. Se as causas das distorções são perfeitamente conhecidas é possível, em princípio, aplicar a transformação inversa para remover essas distorções. Em outros casos, o modelo da distorção não é completamente disponível e são então utilizados os chamados pontos de controle, que são pontos característicos (por exemplo: intersecções de estradas em imagens aéreas) cuja localização é perfeitamente conhecida. Modela-se, então, a transformação por polinômios cujos coeficientes são determinados por sistemas de equações lineares. Uma vez estabelecido o mapeamento inverso da imagem corrigida para a imagem distorcida, resta o problema de se interpolar, na imagem distorcida, os valores de luminância no interior da grade correspondente aos valores inteiros na grade da imagem corrigida. Vários métodos de interpolação para este processo (denominado reamostragem) têm sido propostos. O processo de se determinar pontos de controle correspondentes nas duas imagens pode ser feito de maneira automática utilizando técnicas de registro de imagens, a serem abordadas na secção de análise de imagens.

1.4 - RESTAURAÇÃO E RECONSTRUÇÃO

Enquanto as técnicas de realce de imagens têm geralmente um caráter *ad hoc*, pelo fato de se basearem num critério subjetivo de avaliação, os processos de restauração e reconstrução de imagens podem ter uma formalização matemática mais rigorosa, pois se baseiam em critérios de avaliação objetivos. Contudo, se esses problemas admitem uma solução formal, sua implementação em computador pode ser bastante dificultada pelo esforço computacional exigido ao lidar com problemas bidimensionais. O problema de restauração de imagens refere-se à tarefa de estimar a ima

gem original que sofreu um processo de degradação, envolvendo espalhamento da luz e contaminação por ruído. Esses problemas ocorrem em imagens tiradas próximas ao limite de resolução de instrumentos (Astronomia, Microscopia Eletrônica, imageamento por satélites militares de alta resolução ou mesmo Radiografia Médica) ou ainda em condições adversas (movimento ou turbulência).

Se a imagem original sofreu um processo de "borramento", é possível modelar tal processo através de uma equação de convolução, supondo as condições ideais de linearidade e invariância no espaço. Assim sendo, tomando-se a Transformada de Fourier, verifica-se que é possível, em princípio, recuperar a imagem original pela razão entre as Transformadas de Fourier da imagem borrada e da função de espalhamento. Surgem dois problemas com a utilização desse método:

- 1) A transformada de Fourier da função de espalhamento (função de transferência) pode ter zeros, e o filtro inverso não seria realizável, havendo necessidade de fazer uma realização aproximada.
- 2) A inevitável presença do ruído faz com que o resultado fique completamente mascarado nas frequências espaciais nas quais o ruído, que geralmente contém componentes de alta frequência, predomina. Pode-se aliviar o problema pelo truncamento do filtro a frequências mais baixas.

Tal situação está intimamente ligada ao fato de a chamada equação integral de convolução, que caracteriza o borramento da imagem, ser essencialmente um problema mal condicionado, no sentido de que pequenas perturbações na imagem borrada podem provocar enormes variações na imagem restaurada. A única solução para este tipo de problema está no uso de conhecimento a priori sobre a imagem original.

Este conhecimento pode ser de dois tipos: de determinístico ou estatístico. No primeiro caso, podem ser utilizados os chamados métodos de regularização, onde se minimiza uma expressão quadrática que mede a suavidade da solução. Pode-se mostrar que a solução obtida é equivalente àquela que seria obtida pela utilização de conhecimento estatístico na formulação de Wiener, descrita adiante. Deve-se notar, aqui, que o problema de restauração de imagens é basicamente um problema de deconvolução, que aparece frequentemente nas ciências físicas. No caso de imagens, o problema se torna mais difícil pela enorme carga computacional que pode advir do fato de se tratar de uma situação bidimensional.

No caso de a função de espalhamento ser invariante no espaço, é possível reduzir substancialmente essa carga pelo uso da Transformada Rápida de Fourier ("FFT-Fast Fourier Transform"). Essa Transformada, ao lado de outros algoritmos de rápida execução, como

Transformadas de Walsh-Hadamard, do cosseno e "slant", desempenha um importante papel em processamento de imagens ou de sinais, de modo geral. Outras aplicações da Transformada Rápida de Fourier e de suas congêneres serão examinadas na Seção 1.5 sobre codificação.

A possibilidade de usar conhecimento determinístico a priori para combater o mal condicionado também pode ser explorada na forma de restrições de desigualdade. Uma restrição natural que ocorre em imagens reside no fato de valores de luminância serem necessariamente não-negativos. Pela utilização do critério de mínimos quadrados, pode-se mostrar que daí resulta um problema de programação quadrática. Todavia, o esforço computacional exigido só permite lidar com imagens de pequena resolução ou em áreas localizadas.

O conhecimento a priori do tipo estatístico pode ser incorporado através dos dois primeiros momentos da imagem original e do ruído. Nessas condições, a estimação ótima é dada pelo filtro de Wiener, que também pode ser implementado de maneira eficiente usando a Transformada Rápida de Fourier. Tem havido recentemente tentativas de implementar métodos de restauração não-linear, levando em conta, por exemplo, as características do processo de emulsão fotográfica.

O problema de reconstruir uma imagem por suas projeções tem atraído muita atenção nos últimos anos, particularmente devido à sua aplicação na área de tomografia computadorizada. Esse problema tem uma estrutura matemática semelhante à da restauração de uma imagem. Nos dois casos, as observações resultam de uma combinação linear dos valores da imagem original, além da presença inevitável do ruído.

Assim sendo, um dos tipos de abordagem tem características próximas às dos métodos de restauração; são as chamadas técnicas algébricas, equivalentes a resolver um grande sistema de equações lineares. Como no caso de restauração, elas podem ser encaradas de um ponto de vista determinístico ou estatístico. Esta última formulação pode levar a esquemas recursivos do tipo filtragem de Kalman, onde a reconstrução é atualizada a cada novo conjunto de medidas tomadas.

A possibilidade de usar as ferramentas da Transformada Rápida de Fourier para reconstrução advém do seguinte resultado: dada uma imagem f com uma Transformada de Fourier bidimensional F , se $f\theta$ é a projeção de f segundo uma direção θ , o que foi obtido somando os valores da imagem ao longo dessa direção, então a Transformada de Fourier unidimensional dessa projeção é um corte de F através da origem do plano de frequências na direção $\theta + \pi/2$. Deste modo, pode-se obter uma aproximação de F por várias projeções e reconstruir a imagem pela Transformada Inversa de Fourier.

1.5 - CODIFICAÇÃO

Um dos objetivos dos processos de codificação de imagem é reduzir a quantidade de informação necessária para representar uma imagem com a finalidade de diminuir o tempo, a largura de faixa ou a potência necessários para transmitir a imagem, ou o volume de memória para armazená-la. Este artigo apresenta, a seguir, um breve resumo das principais técnicas de codificação utilizadas em imagens.

As técnicas de codificação podem ser exatas, no sentido de possibilitarem uma exata reconstrução da imagem original, ou aproximadas, que causam alguma distorção, muitas vezes quase imperceptível ao olho humano.

Dentre as técnicas exatas, podem ser mencionadas:

- a) código de Huffman: se os níveis de cinza da imagem não são distribuídos uniformemente, é possível associar uma palavra de código mais curta aos níveis mais frequentes, de modo a reduzir o número médio de bits necessário para representar a imagem;
- b) código de comprimento de corrida: se os níveis de cinza ao longo de uma linha da imagem tendem a ser constantes, por um comprimento suficientemente longo, é possível utilizar este fato, caracterizando a imagem pelos comprimentos dessas corridas e pelos níveis de cinza a elas associados;
- c) código de contornos: se uma imagem é constituída por regiões homogêneas conectadas, uma representação econômica consiste em codificar os contornos das regiões e os níveis de cinza de cada uma delas.

Uma das possibilidades de aproximar a representação de uma imagem consiste em tirar proveito das limitações da visão humana. Assim, em regiões onde o nível de cinza varia lentamente, a amostragem pode ser mais espaçada, o mesmo ocorrendo com a quantização em regiões onde a flutuação é rápida. Esse método pode ser implementado tanto no domínio do espaço quanto no domínio da frequência, mostrando espaçadamente as componentes da Transformada de Fourier da imagem em baixas frequências, o mesmo sendo feito com relação à quantização das frequências altas.

A redundância existente devida à correlação entre os "pixels" adjacentes da imagem pode ser explorada pelas chamadas técnicas preditivas, onde é codificado o sinal resultante da diferença entre "pixels" adjacentes, ou da diferença entre o "pixel" e a predição dele baseada nos "pixels" adjacentes na mesma linha (predição unidimensional), e em linhas adjacentes (predição bidimensional). A possibilidade de redução do número de bits decorre da distribuição fortemente não-uniforme do sinal-diferença (apresentando um pico

em valores baixos), permitindo o uso do código de Huffman. Além disso, as grandes amplitudes tendem a variar rapidamente, o que é aproveitado por uma quantização espaçada. As técnicas diferenciais podem envolver apenas 1 bit na quantização do sinal-diferença (modulação delta), ou mais de 1 bit (DPCM - "differential pulse code modulation"). Estas técnicas são de implementação simples, mas apresentam o problema de propagação de eventuais erros na transmissão da imagem.

O uso de transformadas de imagens tem sido bastante explorado para se efetuar uma compressão de dados pictóricos. As Transformadas de Fourier, Walsh-Hadamard, cosseno e "slant" têm oferecido reduções de 10:1 ou mais na taxa de transmissão. Outra transformação possível, que tem propriedades de otimalidade, é a de Karhunen-Loève, mas tem a desvantagem de não possuir uma implementação por meio de algoritmos rápidos, como é o caso das transformações anteriores, com exceção do caso em que a imagem possui uma função de correlação separável nas direções horizontal e vertical e de tipo markoviana.

Nesse tipo de codificação, o fato de a transformação da imagem ser compressível é fundamental. Por exemplo, na Transformada de Fourier, os coeficientes de frequências mais baixas têm, em geral, grandes amplitudes, e o oposto ocorre com as frequências altas, de modo que os coeficientes devem ser quantizados conforme sua amplitude. Estas observações estão de acordo com os métodos aproximados de digitalização para compressão, mencionados anteriormente. As técnicas de codificação por transformadas, embora de implementação mais complexa que as técnicas de diferença, têm a vantagem de distribuírem efeitos de erros de transmissão, tornando-os menos visíveis. Deve-se mencionar, nesse ponto, que técnicas híbridas, empregando métodos diferenciais e de transformadas, também têm sido desenvolvidas.

As técnicas de codificação podem ser estendidas a imagens com características especiais, como imagens binárias de texto e gráficos, imagens a cores ou multiespectrais, ou ainda, sucessivos quadros de imagens de TV, onde uma pequena fração do número de pontos sofre mudança (codificação entre quadros) de quadro para quadro.

Os limites dos métodos de codificação são estabelecidos pela teoria de razão de distorção. Todavia, além de não especificar como projetar os métodos de codificação para atingir esses limites, essa teoria sofre das deficiências no modelamento estatístico realístico da fonte de sinal (imagens) e no estabelecimento de um critério de fidelidade de que, simultaneamente, seja tratável do ponto de vista matemático e leve em conta as características complexas do sistema visual humano.

1.6 - ANÁLISE DE IMAGENS

A área de análise de imagens engloba várias subáreas mais ou menos independentes, que se rão tratadas brevemente nesta secção.

Algoritmos para detectar bordas em imagens podem ser uma etapa de pré-processamento em vários sistemas de processamento de imagens e reconhecimento de padrões. São largamente utilizadas as técnicas de diferenciação do tipo gradiente ou laplaciano em suas aproximações digitais, seguidas da comparação com um limiar para decidir a respeito da presença ou não de uma borda local. O problema principal reside na susceptibilidade ao ruído inerente a esses operadores. Para contornar o problema, é possível lançar mão de técnicas estatísticas, como filtragem de Wiener, para estimar o laplaciano sob ruído, ou de testes de hipóteses para decidir sobre a presença ou não de algum tipo de borda.

Após a aplicação de operadores locais, poderá ser necessário unir os pontos resultantes de modo a formar uma curva contínua. Vários métodos têm sido propostos para resolver este problema dos quais um dos mais simples envolve a utilização de aproximações poligonais ou polinomiais.

O problema de binarizar uma imagem de tons de cinza ocorre em várias aplicações. Sensores visuais em robôs trabalham muitas vezes com imagens binárias. O procedimento mais natural baseia-se no estabelecimento de um limiar nos níveis de cinza, mas podem surgir problemas quando se deseja detectar um objeto cujos tons não diferem substancialmente do fundo. Nesse caso, o uso de propriedades locais como intensidade média ou de limiares dependentes da localização pode ser útil.

O problema de descrever a forma de objetos em imagens tem importância considerável em áreas como aplicações industriais. As técnicas desenvolvidas podem ser de dois tipos:

- a) descrição de regiões, envolvendo análise dos eixos principais dos objetos ou propriedades geométricas como perímetro e área ou ainda descritores topológicos como o número de Euler, etc.
- b) descrição de periferia como código de cadeias, representações por grafos, descritores de Fourier, etc.

Inúmeras técnicas de análise de textura têm sido desenvolvidas. O problema é relevante em imagens biomédicas ou de sensoriamento remoto, por exemplo. Uma caracterização natural é feita através do espectro de Fourier, mas os resultados até agora não têm sido inteiramente satisfatórios. Outra possível caracterização pode ser feita através das chamadas matrizes de co-ocorrência, que envolvem os histogramas bidimensionais dos níveis de cinza para um dado ângulo e uma dada distância de separação entre os "pixels". Entre

outros esquemas, poderiam também ser citados aqueles que envolvem a densidade de bordas por unidade de área, ou os modelos de séries temporais sazonais.

O registro preciso de duas imagens tiradas em condições diferentes pode ser necessário em aplicações que vão desde imagens aerospaciais de sensoriamento remoto ou meteorológica até imagens biomédicas. A primeira tentativa para resolver este problema consistiu em efetuar uma correlação entre as duas imagens, localizando o ponto de máximo. As dificuldades do método residem, principalmente, na imprecisão da determinação do ponto ótimo e no enorme esforço computacional, a despeito do uso da Transformada Rápida de Fourier. Os métodos mais bem sucedidos têm se baseado em esquemas sequenciais, onde a soma do módulo das diferenças entre "pixels" correspondentes para cada posição candidata a registro é comparada com um limiar pré-terminado, de modo a truncar o processo acumulativo rapidamente. É possível reduzir ainda mais a carga computacional pelo uso de imagens de resoluções diferentes, num esquema hierárquico.

Para finalizar esta secção, podem ser mencionadas as técnicas de propagação e contração de regiões, que permitem simplificar as formas, eliminando lacunas ou pontos isolados. Estes procedimentos têm sido usados em detecção de falhas em inspeção industrial, delimitamento de bordas de textura ou análise de células biológicas.

2 - CLASSIFICAÇÃO DE PADRÕES

2.1 - INTRODUÇÃO

A área de reconhecimento de padrões não está restrita à classificação e descrição de imagens, mas a maior parte do trabalho desenvolvido tem se voltado para aplicações pictóricas. Assim sendo, têm havido tentativas sérias de automatizar as tarefas de leitura de caracteres alfanuméricos; análise de imagens aeroespaciais, biomédicas e de câmaras de bolha em física de alta energia. Destacam-se também aplicações não-pictóricas como reconhecimento de voz e do orador, análise de sinais de sonar, eletrocardiograma, eletroencefalograma e em sismologia, assim como classificação de dados obtidos em antropologia ou zoologia, etc.

O reconhecimento de padrões tem se desenvolvido, em geral, ao redor de duas grandes linhas de abordagem:

- 1) Classificação de padrões, onde o objetivo é atribuir um objeto a uma das possíveis classes. A ferramenta básica é a teoria da decisão estatística, embora várias técnicas empregadas sejam essencialmente determinísticas.
- 2) Reconhecimento sintático (ou estrutural) de padrões, onde se procura uma descrição do objeto em termos de relações entre suas

partes, utilizando a teoria de linguagens formais. Este trabalho vai abordar, em sua segunda parte, apenas a área de classificação de padrões.

As motivações para o desenvolvimento de classificação de padrões podem ser encontradas:

- 1) No desafio de reproduzir como computador as habilidades de tomar decisões que a mente humana possui.
- 2) Na necessidade de desenvolver métodos práticos e eficientes que substituam o homem em vários tipos de tarefas. Algumas dessas tentativas têm revelado grande êxito em ambientes controlados, como é o caso de reconhecimento de caracteres impressos de cheques bancários. Grandes esforços de pesquisa têm sido feitos nas situações em que este controle é menos rígido, como, por exemplo, no reconhecimento de caracteres manuscritos.

Os conhecimentos atuais de neurofisiologia não têm sido suficientes para reproduzir, na máquina, as funções de percepção de seres muito primitivos. Todavia, mesmo diante dessa insuficiência, tem havido êxito na resolução de diversos problemas de reconhecimento de padrões, particularmente nas tarefas mais simples de classificação, ou seja, a de atribuir um objeto a uma de diversas classes. Os métodos desenvolvidos têm tido por base, em geral, uma formulação estatística que pode estar distante dos processos perceptuais em seres vivos.

Deste modo, é feito, inicialmente, um conjunto de medidas por um sensor. Não serão discutidos nesta seção os problemas referentes às medidas que devem ser tomadas nem a maneira como deve ser projetado o sensor, uma vez que tais problemas são fortemente dependentes da aplicação. Os dados obtidos são representados por meio de um vetor de medidas num espaço de uma determinada dimensão. Normalmente, é necessário efetuar um processo de redução da dimensionalidade (seleção de atributos) antes da fase de classificação, de modo a reduzir o esforço computacional nesta última fase.

A fase de classificação se reduz essencialmente a uma partição do espaço de atributos, de tal maneira que, idealmente, os pontos pertencentes a uma determinada classe caiam sempre na partição correspondente. Como esse objetivo ideal nem sempre pode ser alcançado, deve-se estabelecer a partição minimizando algum critério que, de acordo com a formulação estatística do problema, é a probabilidade de erro ou um risco médio, dependendo de custos apropriadamente escolhidos. Quando as classes apresentam pequena variação em torno de um protótipo, a partição pode ser efetuada de modo a atribuir o objeto ao protótipo mais próximo. Contudo, em muitas situações, a variação é apreciável, e técnicas mais sofisticadas de decisão estatística devem ser empregadas.

O problema de classificação de padrões está intimamente ligado ao problema de aprendizagem do estatístico, uma vez que nas situações reais é raro ter uma descrição estatística completa das densidades de probabilidade que caracterizam cada classe.

2.2 - EXTRAÇÃO DE ATRIBUTOS

As técnicas de extração de atributos podem ser consideradas como uma forma de compressão de dados, devido ao caráter de redução de dimensionalidade que elas possuem. As técnicas de extração de atributos dividem-se em duas categorias:

a) Transformações no Espaço de Atributos

O objetivo aqui é transformar o espaço de atributos original em um espaço menor dimensionalidade, preservando ao máximo a representação dos padrões e/ou a discriminação entre as classes. No primeiro caso têm sido usados critérios de mínimo erro médio quadrático ou entropia. A Transformação de Karhunen-Loève, envolvendo os vetores e os valores próprios da matriz de covariância desempenha um importante papel nesta área. Quanto à discriminação entre classes, é possível utilizar transformações lineares ou não-lineares, sendo obtido um resultado superior no último caso, ao preço, todavia, de uma carga computacional que pode ser bastante superior. Frequentemente é necessário usar técnicas iterativas ou interativas, através de um "display".

Pode haver conflito entre os critérios de representação e os de separabilidade entre classes, e seria desejável que se estabelecesse um compromisso entre esses fatores na escolha da transformação ótima.

b) Medidas de Informação e Distância

Em última análise, o critério a ser utilizado para avaliar o desempenho de um sistema de classificação deveria ser a probabilidade de erro, mas existem grandes dificuldades em empregar esse critério diretamente na redução da dimensionalidade. Assim sendo, utilizam-se critérios indiretos expressos por medidas de distância entre distribuições, que fornecem limites superiores e/ou inferiores para a probabilidade de erro. Várias medidas têm sido propostas, e investigadas as relações entre elas. A divergência e a distância de Bhattacharyya têm a vantagem de serem mais fáceis de computar do que outras medidas propostas.

2.3 - MÉTODOS DE CLASSIFICAÇÃO

O problema central de classificação, para um dado vetor de atributos observado, envolve o cálculo dos valores das funções densidade de probabilidade, relativas às várias classes, e a determinação do máximo valor, a cuja classe é atribuído o referido vetor. Quando as probabilidades a priori são conhecidas e os custos de decisão são escolhidos,

o problema se enquadra na chamada teoria de decisão bayesiana, ao passo que se as classes são equiprováveis, tem-se a decisão por máxima verossimilhança.

A possibilidade de utilizar os esquemas de testes sequenciais de hipóteses tem sido explorada em classificação de padrões, dada a limitação por vezes existentes no número de amostras disponíveis.

Outra idéia que pode ser utilizada em alguns casos é a de utilizar informação de correlação entre amostras sucessivas (como "pixels" adjacentes em imagens) para melhorar a classificação. Esquemas desse tipo estão incluídos na chamada teoria de decisão composta.

Os demais métodos de classificação descritos a seguir referem-se essencialmente ao problema que ocorre quando essas densidades de probabilidade não são conhecidas completamente.

Quando a forma funcional das densidades é conhecida, exceto por seus parâmetros, daí resulta um problema de aprendizado, que é um processo de estimação dos parâmetros desconhecidos. Duas situações diferentes podem ocorrer, dependendo do tipo de amostras disponíveis:

- a) amostras de classificação conhecida;
- b) amostras de classificação não-conhecida;

No primeiro caso, tem-se o chamado aprendizado com supervisão, e no segundo, o aprendizado sem supervisão.

No aprendizado com supervisão, a disponibilidade de um rótulo nas amostras simplifica consideravelmente o esforço computacional. Se os parâmetros são considerados não-aleatórios, as técnicas de estimação por máxima verossimilhança podem ser empregadas. Já na situação em que os parâmetros são considerados variáveis aleatórias, cai-se num esquema bayesiano. Como os estimadores dos parâmetros contêm erros, o desempenho do classificador não é tão bom como no caso em que os parâmetros são conhecidos. Todavia, com a disponibilidade de um grande número de amostras, a qualidade dos estimadores melhora, assim como o desempenho do classificador.

Outra situação que ocorre frequentemente é quando a forma das densidades de probabilidade não é conhecida. As formas paramétricas usuais raramente descrevem com fidelidade as densidades encontradas na prática. Em geral, essas formas paramétricas têm a forma unimodal, enquanto os histogramas reais são muitas vezes multimodais. Deste modo, recorre-se às chamadas técnicas não-paramétricas. Algumas dessas técnicas envolvem a estimação das densidades condicionais de cada classe a partir das amostras. Outras técnicas estimam diretamente as probabilidades a posteriori das classes, que são usadas como fun-

ções discriminantes, sem passar pelas densidades condicionais. Esses procedimentos estão relacionados com o método de classificação de uma amostra pelo protótipo de classificação conhecida mais próximo (vizinho mais próximo). Em geral, as técnicas não-paramétricas trocam a necessidade da forma funcional das densidades ser conhecida por uma necessidade de um número maior de amostras.

Suponha-se agora que sejam oferecidas ao classificador, apenas amostras de classificação conhecida. Nessas condições, o classificador deve ajustar iterativamente as superfícies de decisão, de modo a, se possível, separar perfeitamente as amostras. Esses procedimentos, que tiveram muita popularidade no início da década de 60, são os algoritmos do tipo "perceptron" ou de relaxação. Para lidar com a situação em que as classes não são separáveis linearmente, foram desenvolvidos os chamados procedimentos de mínimo erro médio quadrático, que convergem ainda na situação não-separável. A função objetivo que caracteriza os algoritmos do tipo "perceptron" também pode ser minimizada por métodos de programação linear. Devem ser mencionadas também, entre as funções discriminantes lineares, as chamadas funções de potencial, cujo desenvolvimento se inspirou nos métodos da Física Clássica.

Se a única informação possível para projetar o classificador consiste em amostras não-rotuladas, recorre-se às técnicas de agregação de dados ("clustering"). Essas técnicas tiveram origem nos processos de taxonomia numérica das Ciências Biológicas e, com o advento do computador digital, passaram a ser largamente utilizadas. Em geral, os métodos de agregação são não-supervisionados e até mesmo o número de classes pode não ser determinado a priori. As técnicas desenvolvidas baseiam-se frequentemente na minimização de um critério derivado de uma medida de similaridade entre amostras. Esse critério pode ser de mínima variância, envolvendo as matrizes de espalhamento entre classes ou dentro de classes. Técnicas baseadas em teoria de grafos, análise combinatória ou ainda na teoria de conjuntos nebulosos ("fuzzy sets") têm sido desenvolvidas. Os procedimentos de agregação podem ser classificados como hierárquicos ou não-hierárquicos. Nos procedimentos hierárquicos, as partições das amostras são feitas em níveis, por aglomeração, ou divisão. A característica hierárquica se baseia na propriedade de que se duas amostras estão no mesmo agregamento num nível do processo, elas assim permanecerão até o fim do mesmo. Os exemplos mais conhecidos de agregação hierárquica ocorrem em Biologia, onde indivíduos se agrupam em espécies, espécies em gêneros; gêneros em famílias, etc. A despeito de sua enorme importância prática, os métodos de agregação frequentemente se ressentem de uma base mais sólida nas decisões de dividir ou aglomerar amostras, ou na escolha do número de aglomeramentos. Assim sendo, desenvolve-se atualmente um esforço considerável no sentido de validar estatística

mente esses processos.

É usual em problemas de classificação de padrões assumir que, na fase de aprendizado dos parâmetros, estes sejam estimados com um número ilimitado de amostras. Todavia, isso nem sempre corresponde à realidade. A frequente disponibilidade de apenas um pequeno número de amostras pode afetar bastante o desempenho do classificador. É uma área ativa de pesquisa o estudo do relacionamento que deve existir entre a dimensionalidade das amostras e seu número, e de como isto afeta a tarefa de classificação.

Em última análise, o desempenho do classificador é medido pela probabilidade de erro de classificação. O procedimento de avaliar essa probabilidade pela classificação do próprio conjunto de amostras de treinamento é considerado otimista, de modo que é necessário dividir o conjunto de amostras numa parte de treinamento e numa parte de teste. É preferível fazer a estimativa da probabilidade de erro em N passagens, em cada uma das quais uma observação é mantida para teste, e as restantes $N-1$ são usadas para treinamento.

2.4 - CONCLUSÕES

Depois de cerca de duas décadas de intenso desenvolvimento, as áreas de Processamento de Imagens e Reconhecimento de Padrões apresentam ainda importantes e inúmeros problemas a serem resolvidos. Particularmente na segunda área, algumas das promessas iniciais ainda não foram cumpridas, tendo-se verificado que certos tipos de problemas de percepção são, na realidade, muito mais complexos do que a princípio pareciam. Entretanto, o desenvolvimento, tanto do ponto de vista teórico como prático, já tem sido suficiente para abrir largos horizontes às aplicações de alcance cada vez maior.

BIBLIOGRAFIA (Secções 1 e 2)

- AGRAWALA, A.K., ed. *Machine recognition of patterns*. New York, N.Y., IEEE Press, 1976.
- ANDREWS, H.C. *Introduction to mathematical techniques in pattern recognition*. New York, N.Y., John Wiley, 1972.
- BERNSTEIN, R., ed. *Digital image processing for remote sensing*. New York, N.Y., IEEE Press, 1978.
- COMPUTER GRAPHICS AND IMAGE PROCESSING. New York, Academic, 1972-bim.
- DUDA, R.O.; HART, P.E. *Pattern classification and scene analysis*. New York, N.Y., John Wiley, 1973.
- FU, K.S.; ROSENFELD, A. *Pattern recognition and image processing*. *IEEE Transactions on Computers*, C-25(12): 1336-1346, Dec. 1976.

FUKUNAGA, K. *Introduction to statistical pattern recognition*. New York, N.Y., Academic, 1972.

GONZALEZ, R.C.; WINTZ, P. *Digital image processing*. Reading, MA., Addison Wesley, 1977.

HALL, E.L. *Computer image processing and recognition*. New York, N.Y., Academic, 1979.

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. New York, N.Y., IEEE Computer Society, 1979- , bim.

PATTERN RECOGNITION. Oxford, Pergamon, 1968- , bim.

PRATT, W.K. *Digital image processing*. New York, N.Y., John Wiley, 1978.

ROSENFELD, A. *Image processing and recognition*. In.: YOVITZ, M.C. *Advances in computers*. New York, N.Y., Academic, 1979. V. 18, Cap. 1, p. 1-57.

ROSENFELD, A.; KAK, A.C. *Digital picture processing*, 2ª ed., New York, N.Y., Academic, 1982.

TOU, J.T.; GONZALEZ, R.C. *Pattern recognition principles*. Reading, MA., Addison Wesley, 1974.

YOUNG, T.Y.; CALVERT, T.W. *Classification, estimation and pattern recognition*. New York, N.Y., American Elsevier, 1974.

3 - VISÃO ROBÓTICA

3.1 - INTRODUÇÃO

Ultimamente vem crescendo o uso de robôs, na indústria, em tarefas cada vez mais complexas. As razões apontadas para isto tem sido a existência, no processo de manufatura, de tarefas insalubres ou perigosas e de tarefas repetitivas e monótonas onde o desempenho humano tende a ser baixo. Some-se a isto a pressão para baixar os custos a fim de manter a competitividade, sobretudo no mercado internacional.

Pode-se definir robô como sendo um "manipulador programável capaz de desempenhar trabalho útil, automaticamente, sem a ajuda humana" (Allan, 1979). Embora esta definição seja incompleta, por excluir tarefas de inspeção que não envolvam manipulação, ela ressalta uma característica importante que é a programabilidade do robô. Por exemplo, um robô programado para soldar ou pintar (tarefas típicas) um tipo de automóvel pode ser reprogramado para outros tipos, caso necessário.

Mais recentemente, os robôs têm sido dotados de sensores óticos ou táteis. Estes sensores visam permitir uma maior complexidade nas tarefas executadas, assim como aumentar sua flexibilidade. Um robô dotado de "visão" pode dispensar que as peças manipuladas venham em posições rigidamente predeterminadas e presas a estas posições. Isto poupa o traba-

lho de posicionar as peças com precisão e economiza esteiras especiais. Além disso, algumas tarefas são especialmente adaptadas à visão robótica como a inspeção e o controle de qualidade de placas de circuitos impressos. A robotização do controle de qualidade permite que *todas* as peças produzidas sejam examinadas, em vez de apenas uma pequena amostra, como é normalmente feito.

Robôs com "visão" são usados em tarefas que vão desde a simples inspeção até complexas tarefas de montagem que exigem um alto grau de interação entre o sistema de visão e o braço manipulador. A introdução de visão a robôs não tem sido maior devido, principalmente, ao custo - um robô "inteligente" que custaria US\$ 60000, com visão passaria a US\$ 100000 (Allan, 1979). Contudo, a perspectiva é de que os sistemas de visão para robôs caiam de preço.

Esta seção procura dar uma idéia do que é visão robótica, suas áreas de aplicação e a natureza dos algoritmos e técnicas usadas. Sempre que possível, os conceitos são ilustrados com exemplos de sistemas existentes. Atualmente um grande número de publicações apresenta trabalhos sobre visão por computador e, mais particularmente, visão robótica. Os números de maio de 1980 e de dezembro de 1982 da revista "Computer" foram dedicados à percepção por computador e robótica, respectivamente. Outras fontes de caráter geral também utilizadas nesta seção foram as de Barrow e Tenenbaum (1981), Chin e Harlow (1982) e Kruger e Thompson (1981).

3.2 - CONCEITOS GERAIS

A primeira pergunta que surge é "o que é visão robótica?", ou seja, quais componentes materiais ("hardware") e quais os algoritmos e procedimentos ("software") que caracterizam a visão robótica. Inicialmente deve-se dizer que visão robótica é um caso particular de visão por computador. Um sistema computacional, para ser considerado capaz de visão, deve satisfazer os seguintes critérios (Agin, 1980):

- 1) O sistema deve processar imagens bidimensionais; o modo como estas imagens são obtidas varia grandemente de sistema para sistema.
- 2) O processamento feito nas imagens deve extrair, explicitamente, informação útil das imagens; exemplos de informação útil, no caso de robôs, são a identidade de peças, suas localizações, etc.

Na definição está implícita a existência de um computador programável. A definição exclui sistemas baseados em sensores óticos isolados como, por exemplo, fotocélulas, e os que somente manipulam imagens como, por exemplo, sistemas para realçamento ou armazenamento de imagens. Estes sistemas, todavia, podem constituir-se em subsistemas importantes de sistemas de visão por computador.

Quando o sistema de visão por computador é empregado (ou pode ser aplicável) numa operação de manufatura, pode ser caracterizado como de visão robótica. Esta restrição tem consequências importantes tanto na natureza dos componentes materiais, quanto nos programas que farão o processamento das imagens. Um sistema de visão robótica deve ser robusto para suportar as condições de trabalho num ambiente industrial e operar continuamente com poucas falhas para minimizar os custos de manutenção. Deve ser também de baixo custo para que compense, economicamente, sua introdução em linhas de produção. Além disso, existem padrões mínimos de desempenho que estes sistemas devem satisfazer. Estes padrões mínimos dependem, naturalmente, do campo de aplicação; um segundo por imagem constitui um limite superior, na maioria das operações de manufatura (Agin, 1980).

A necessidade de robustez e baixo custo implica, geralmente, uma tendência para a simplicidade dos sistemas. Por outro lado, as restrições de desempenho - efetividade e velocidade (o sistema deve executar *bem* sua tarefa num dado intervalo de tempo) - implicam sistemas sofisticados, material e logicamente. Estas duas tendências são contraditórias e boa parte dos esforços consiste em achar o ponto ótimo de custo e efetividade.

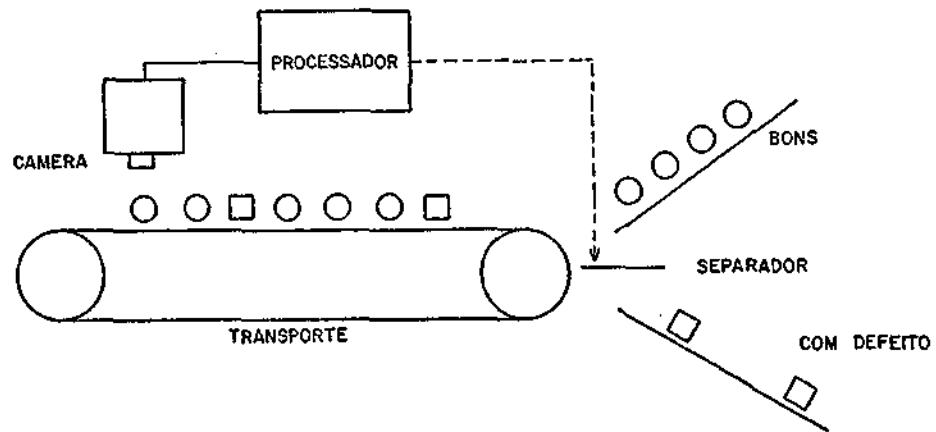
Para diminuir a complexidade da tarefa de visão e aumentar sua velocidade, duas técnicas frequentemente utilizadas são:

- 1) controlar o ambiente através, por exemplo, de iluminação especial;
- 2) especializar os sistemas para "classes de aplicações".

Quanto à segunda técnica, pode-se dizer que ainda se está longe de ter uma visão robótica de propósito geral. Não há, inclusive, consenso entre os pesquisadores da área de que tal sistema seja possível (Jarvis, 1982). Deste modo, os sistemas de visão robótica costumam ser divididos nas seguintes áreas de aplicação:

- 1) inspeção,
- 2) manuseio de peças
- 3) montagem.

Na inspeção, a peça é examinada para a verificação de defeitos. Um exemplo de inspeção é o controle de qualidade de placas de circuito impresso (Jarvis, 1980; Baird, 1978; Chin and Harlow, 1975). A verificação se uma peça foi corretamente montada também pode ser considerada como uma tarefa de inspeção. A figura 1 mostra, esquematicamente, o diagrama de um sistema automático de inspeção visual.

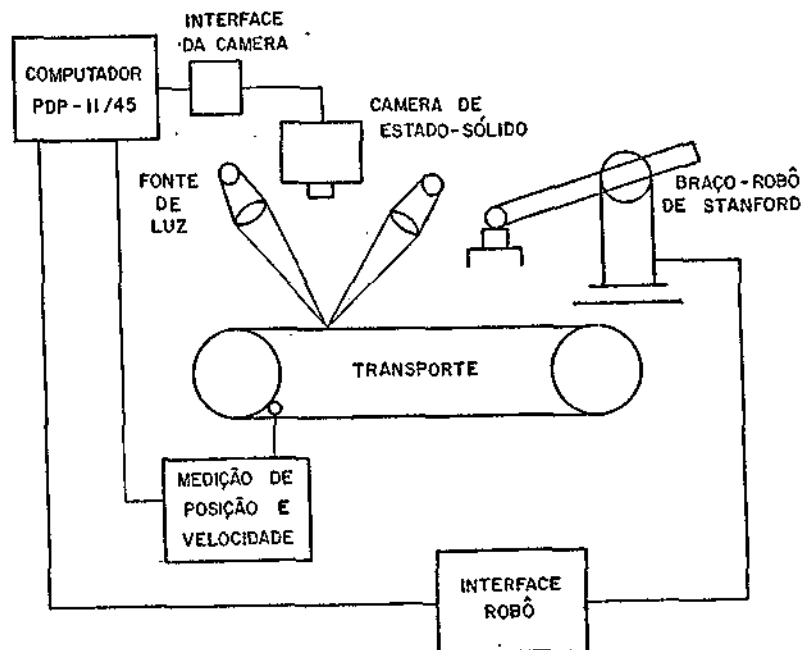


"FIG. 1" - DIAGRAMA DE UM SISTEMA AUTOMÁTICO DE INSPEÇÃO VISUAL

FONTE: Chin and Harlow (1982)

Operações comuns de manuseio de peças são do tipo "cata e posiciona" ("pick and place"). Neste caso, o processamento envolve a identificação de peças, bem como a determinação de suas posições. A interação da visão com o braço manipulador é bem maior que nas tarefas de inspeção. Esta última característica introduz um nível a mais de dificuldade. Exemplo de tal sistema é o CONSIGHT-I da General Motors, capaz de localizar, identificar, catar e posicionar peças colocadas numa esteira rolante. A Figura 2 mostra, simplificada, o diagrama geral do sistema CONSIGHT-I.

As tarefas de montagem são as mais complexas de todas, pois incluem a inspeção e o manuseio. Podem exigir a presença de vários robôs como o sistema projetado pela Westinghouse para a montagem de pequenos motores elétricos (Sugarman, 1980), ou vários sensores óticos e táticos associados a vários braços como o sistema usado pela Hitachi para a montagem de aparelhos aspiradores de pó (Kashioka et alii, 1977). A montagem de peças muitas vezes envolve tarefas tais como procurar uma peça, identificá-la, determinar sua posição, orientar a peça em posições de terminadas, guiar a inserção e inspecionar



"FIG. 2" - SISTEMA CONSIGHT-I

FONTE: Gonzalez and Safabakhsh (1982)

resultados intermediários e finais. Em todas estas tarefas a visão pode desempenhar papel importante.

3.3 - PROCESSAMENTO DE IMAGENS EM ROBÔS

O primeiro passo num sistema de visão robótica é converter a informação visual, presente na cena, em sinais elétricos. Isto é feito por meio de sensores óticos. Os sensores óticos mais comumente usados são as câmeras "vidicon" e as matrizes ou vetores de sensores de estado sólido.

A câmera vidicon apresenta a vantagem de baixo custo, uma vez que é o dispositivo usado na geração de sinal de televisão e, portanto, produzida em larga escala. Entretanto sua aplicabilidade é reduzida devido à fragilidade e vida curta do tubo para aplicações industriais.

Câmeras baseadas em dispositivos de estado sólido são compostas de matrizes (ou vetores) de elementos fotossensíveis, que são resistentes e de vida longa. As resoluções para matrizes vão desde 64x64 até 512x512 (esta última comparável com a resolução de televisão) e para vetores até 2048 elementos. No caso de vetor de elementos, a imagem bidimensional é obtida através de uma varredura da cena pelo sensor e de uma amostragem, no tempo, dos sensores.

Os sinais elétricos obtidos são, então transformados em "níveis de cinza" que são inteiros num dado intervalo. Uma distinção importante a fazer na classificação dos sistemas é se o sistema usa dois níveis de cinza (0 e 1) - sistemas binários - ou se usa vários níveis de cinza. A conversão do sinal elétrico em níveis de cinza é mais simples no caso binário, pois basta um circuito de limiarização. No caso de vários níveis, torna-se necessário um conversor analógico-digital. Em geral, o número de níveis de cinza usado é uma potência de 2, sendo comuns os valores 16 e 64.

Os sistemas binários são mais baratos também por necessitarem de menos espaço para armazenamento das imagens. As operações realizadas nas imagens binárias são, na maioria, operações lógicas para as quais os computadores têm instruções específicas, o que torna rápido o processamento. Tais sistemas, contudo, exigem imagens de alto contraste, o que pode, às vezes, ser conseguido com iluminação especial. O limiar de binarização deve ser, também, bem escolhido; alguns sistemas usam limiares variáveis (Ejiri et alii, 1973) que acompanham a variação da média do sinal elétrico.

Sistemas de vários níveis requerem recursos computacionais (processamento e armazenamento) consideravelmente maiores. Contudo, eles são potencialmente mais flexíveis que os sistemas binários; algoritmos mais sofisticados de processamento de imagens também podem ser usados.

Conjectura-se (Jarvis, 1982) que sistemas de propósito geral usarão, necessariamente, vários níveis.

Pode-se dividir em três tipos a maior parte dos algoritmos de processamento de imagens aplicados à visão robótica:

- 1) métodos "ad hoc",
- 2) métodos de casamento direto de formas,
- 3) métodos que envolvem segmentação, descrição e classificação.

Os métodos "ad hoc" não se enquadram em nenhum paradigma de processamento de imagens. Embora muitas vezes sejam os métodos mais efetivos para determinadas aplicações, têm o inconveniente da pouca ou nenhuma flexibilidade, ou seja, valem somente para o problema específico para os quais foram desenvolvidos.

O casamento direto de formas ("template matching") baseia-se na comparação ponto a ponto da imagem, ou parte desta, com um padrão previamente armazenado. Esta comparação pode visar identificar uma peça ou determinar a existência ou não de defeitos. Há várias maneiras de comparar imagens ponto a ponto para a determinação de semelhanças ou dessemelhanças. A operação "equivalência lógica" aplicada a imagens binárias resulta numa imagem que tem valor 1 (um) nos pontos onde as imagens coincidem. Para imagens não-binárias, pode-se usar medidas de correlação entre as imagens.

Na inspeção de placas de circuito impresso, pode-se comparar as imagens obtidas com imagens armazenadas de placas sem defeitos. Este procedimento "direto" é, contudo, raramente usado, dado o esforço computacional necessário. Geralmente são feitas somente comparações locais (pequenas janelas) da imagem com padrões ideais, sem defeitos. Verificou-se que isto era suficiente, pois poucos padrões eram capazes de descrever a maioria das placas sem defeitos (Jarvis, 1980). Este método tem a vantagem de ser menos sensível a erros de alinhamento. Kashioka et alii (1976) usam casamento local de formas para determinar o posicionamento de circuitos para a solda automática. A eficiência alcançada pelo sistema, no qual um computador controla 50 máquinas de solda, é acima de 99%, muito superior ao desempenho humano.

O método mais geral de processamento, usado sobretudo na identificação de peças, é o consistido de três passos:

- 1) segmentação,
- 2) descrição,
- 3) classificação.

Na segmentação, procura-se particionar a imagem em seus objetos constituintes. Estes objetos são, em geral, caracterizados por suas bordas que na imagem correspondem a variações bruscas do nível de cinza, e por constituírem regiões homogêneas na imagem. Os algoritmos de segmentação propostos na literatura, em geral, procuram usar estas características, detetando bordas na imagem ou agrupando pontos vizinhos com níveis de cinza semelhantes. Uma combinação dos dois métodos também é possível.

Uma vez segmentada a imagem, extraem-se características que descrevem os objetos encontrados. Estes atributos podem se referir a forma, posição ou nível de cinza do objeto. Adicionalmente a estes atributos, podem ser extraídos atributos que dependem do relacionamento espacial dos objetos da imagem. A Tabela 1 mostra alguns atributos que podem ser extraídos.

TABELA 1

ALGUNS ATRIBUTOS EXTRAÍVEIS DE OBJETOS NUMA IMAGEM

TIPO	ATRIBUTOS
forma	área, perímetro, relação entre perímetro ao quadrado e área, momentos de inércia, estatísticas de raios que vão do centro do objeto à sua borda, número e área dos buracos, dimensões do retângulo envolvente.
posição	centro, centro do retângulo envolvente, ângulos dos eixos maior e menor do objeto.
nível de cinza	estatísticas dos níveis de cinza dos pontos do objeto (média, desvio padrão)

A classificação dos objetos é feita baseada nos atributos extraídos. Através de uma amostra de treinamento, constituída de objetos cuja identidade é conhecida, determinam-se valores típicos para os atributos considerados. A classificação é feita comparando-se os valores dos atributos dos objetos, cuja identidade se quer determinar, com valores típicos armazenados ou, então, com estatísticas destes valores. Uma regra de decisão bastante usada por sua simplicidade é a do "vizinho mais próximo": a classe do objeto é dada pela classe do objeto da amostra de treinamento, cujos valores estejam mais próximos dos valores dos atributos do objeto que se quer identificar.

O sistema Vision Module, desenvolvido pela SRI International é um bom exemplo de sistema que faz a identificação de peças segundo os passos de segmentação, descrição e classificação (Gleason and Agin, 1979). O sistema Vision Module é composto de uma câmara de

estado-sólido 128x128, uma interface que produz imagens binárias e um microcomputador LSI-11. Alguns atributos extraídos durante a segmentação são: posição do retângulo envolvente, primeiros e segundos momentos de área e perímetro (comprimento e coordenadas). A classificação pode ser feita usando-se a regra do vizinho mais próximo ou através de um procedimento sequencial ("binary decision tree") em que, em cada passo, um único atributo é considerado. O sistema da SRI é bastante flexível, isto advém do fato de ser possível adaptar o sistema para identificar um conjunto diferente de peças simplesmente alterando-se a amostra de treinamento.

3.4 - CONCLUSÕES

Nesta seção procurou-se mostrar a motivação existente para a introdução de visão em robôs industriais, bem como dar uma idéia de como os sistemas de visão robótica são presentemente organizados. Dadas as restrições de custo, muitas vezes torna-se impossível conciliar flexibilidade com eficiência. Mesmo em sistemas que se propõem a generalidade, os algoritmos usados são relativamente simples, pensando as limitações do algoritmo com um controle maior do ambiente.

Mantida a tendência atual de queda dos preços dos circuitos integrados, é de se esperar que sistemas mais poderosos tornem-se viáveis. O processamento de imagens mais complexas, decorrentes de tarefas mais elaboradas ou de ambientes menos controlados, exigirá o uso de técnicas mais sofisticadas como, por exemplo, as usadas em Inteligência Artificial.

Numa das técnicas de Inteligência Artificial para análise de cenas a informação sobre os objetos é armazenada sob a forma de um grafo rotulado, onde os nós representam os objetos ou partes destes, e os ramos representam as relações entre os objetos. A classificação consiste em achar um mapeamento entre a imagem e o grafo, de tal modo que as regiões da imagem possam ser rotuladas com os rótulos dos nós, e as relações espaciais entre as regiões sejam compatíveis com as relações que rotulam os ramos do grafo. Embora este método implique, a princípio, procedimentos muito custosos computacionalmente, é possível, em alguns casos, simplificar a busca do mapeamento ótimo (Barrow and Tenenbaum, 1981).

Um outro problema que tem sido pesquisado é o de extrair informação tridimensional a partir de uma ou várias imagens bidimensionais. Os sistemas atualmente em uso são muito sensíveis a oclusões que possam ocorrer em ângulos das vistas. Estes sistemas estão longe de ter um desempenho semelhante ao humano que usa várias fontes, tais como visão binocular, perspectiva, gradiente de textura, etc., na extração da informação tridimensional.

Os componentes materiais dos sistemas de vi

são robótica são, também, temas de várias pesquisas. A necessidade maior é a de processar grandes quantidades de dados em tempo real ou quase-real. O processamento de imagens é uma área que ganhará muito com o advento de processamento paralelo maciço. Muitos algoritmos de processamento têm a característica de poderem ser executados paralelamente para todos os pontos da imagem em uma ou várias iterações (especialmente os algoritmos ditos visão "de baixo nível" ("low-level vision")). A arquitetura de computadores natural para tais aplicações é a de matrizes bidimensionais de processadores ("cellular arrays") (Rosenfeld, 1983). Atualmente as matrizes celulares 128x128 estão sendo pesquisadas (Rieger, 1981) e não tardará o dia em que matrizes deste porte ou mesmo maiores estarão comercialmente disponíveis.

REFERÊNCIAS BIBLIOGRÁFICAS (Secção 3)

- AGIN, G.J.; "Computer vision systems for industrial inspection and assembly", *Computer*, vol. 13(5), May 1980, pp. 11-20. (Special issue on machine perception).
- ALLAN, R., "Busy robots spur productivity", *IEEE Spectrum*, vol. 17(9), Sept. 1979, pp. 31-36.
- BAIRD, M.L., "SIGHT-I: a computer vision system for automated IC chip manufacture", *IEEE Trans. on Systems, Man and Cybernetics*, vol. SMC-8, Feb. 1978, pp. 133-139.
- BARROW, H.G.; TENENBAUM, J.M., "Computational vision", *Proceedings of the IEEE*, vol. 69(5), May 1981, pp. 572-595.
- CHIN, R.T.; HARLOW, C.A.; "Automated inspection of electronic assemblies", *Computer*, vol. 8(1), Jan. 1975, pp. 36-45.
- CHIN, R.T.; HARLOW, C.A., "Automated visual inspection: a survey", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-4(6), Nov. 1982, pp. 557-573.
- EJIRI, M.; UNO, T.; MESE, M.; IKEDA, S., "A process for detecting defects in complicated patterns", *Computer Graphics and Image Processing*, 1973, pp. 326-339.
- GLEASON, G.J.; AGIN, G.J., "A modular system for sensor-controlled manipulation and inspection", *Proceedings of the 9th International Symposium on Industrial Robots*, Society of Manufacturing Engineers and Robot Institute of America, Washington D.C., Mar. 1979, pp. 57-70. Apud Agin, 1980.
- GONZALEZ, R.C.; SAFABAKHSH, R., "Computer vision techniques for industrial applications and robot control", *Computer*, vol. 15(12), Dec. 1982, pp. 17-32, (Special issue on robotics and automation).
- JARVIS, J.F., "Automated visual inspection of printed wiring boards by local pattern matching", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 2(1), Jan. 1980, pp. 77-82.
- JARVIS, J.F., "Research directions in industrial machine vision: a workshop summary", *Computer*, vol. 15(12), Dec. 1982, pp. 55-61. (Special issue on robotics and automation).
- KASHIOKA, S.; EJIRI, M.; SAKAMOTO, Y., "A transistor wire-bonding system utilizing multiple local pattern matching techniques", *IEEE Trans. on Systems, Man and Cybernetics*, vol. SMC-6(8), Aug. 1976, pp. 562-569.
- KASHIOKA, S.; TAKEDA, S.; SHIMA, Y.; UNO, T.; HAMADA, T., "An approach to the integrated intelligent robot with multiple sensory feedback visual recognition techniques", *Proceedings of the 7th International Symposium on Industrial Robots*, 1977, pp. 531-538. Apud Yashida e Tsuji, 1980.
- KRUGER, R.P.; THOMPSON, W.B., "A technical and economic assesment of computer vision for industrial inspection and robotic assembly", *Proceedings of the IEEE*, vol. 69(12), Dec. 1981, pp. 1524-1538.
- RIEGER, C., "ZMOB: doing it in parallel", *Proceedings of the Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, 1981, pp. 119-124. Apud Rosenfeld, 1983.
- ROSENFELD, A., "Parallel image processing using cellular arrays", *Computer*, vol. 16(1), Jan. 1983, pp. 14-21. (Special issue on computer architecture for image processing).
- SUGARMAN, R., "The blue-collar robot", *IEEE Spectrum*, vol. 17(9), Sept. 1980, pp. 53-57.
- YASHIDA, M.; TSUJI, S., "Industrial computer vision in Japan", *Computer*, vol. 13(5), May 1980, pp. 50-64. (Special issue on machine perception).