

1. Classificação <i>INPE-COM.4/RPE</i> <i>C.D.U.: 621.376.5:621.3.01:519.2</i>		2. Período	4. Distribuição	
3. Palavras Chaves (selecionadas pelo autor)  <i>PROCESSAMENTO DE IMAGENS</i> <i>CLASSIFICAÇÃO DE PADRÕES</i>			interna <input type="checkbox"/>	externa <input checked="" type="checkbox"/>
5. Relatório nº <i>INPE-2090-RPE/320</i>	6. Data <i>Junho, 1981</i>		7. Revisado por <i>Flávio R.D. Velasco</i>	
8. Título e Sub-Título  <i>PROCESSAMENTO DE IMAGENS</i> <i>E CLASSIFICAÇÃO DE PADRÕES</i>			9. Autorizado por  <i>Nelson de Jesus Parada</i> Diretor	
10. Setor <i>DSE/DIN</i>	Código <i>2704</i>		11. Nº de cópias <i>22</i>	
12. Autoria <i>Nelson Delfino d'Ávila Mascarenhas</i>			14. Nº de páginas <i>25</i>	
13. Assinatura Responsável <i>Nelson Mascarenhas</i>			15. Preço	
16. Sumário/Notas  <i>Este artigo apresenta de maneira tutorial um resumo das principais áreas de atividade corrente em Processamento de Imagens e Classificação de Padrões.</i>				
17. Observações <i>Trabalho a ser apresentado no Simpósio sobre "Processamento e Reconhecimento de Imagens por Computador" na 33ª Reunião Anual da SBPC, Salvador, julho de 1981.</i>				

## PROCESSAMENTO DE IMAGENS E CLASSIFICAÇÃO DE PADRÕES

Nelson Delfino d'Ávila Mascarenhas  
Instituto de Pesquisas Espaciais  
Conselho Nacional de Desenvolvimento Científico e Tecnológico

### RESUMO

Este artigo apresenta de maneira tutorial um resumo das principais áreas de atividade corrente em Processamento de Imagens e Classificação de Padrões.

## ÍNDICE

	Pág.
ABSTRACT .....	<i>iv</i>
1. <u>INTRODUÇÃO</u> .....	1
2. <u>PROCESSAMENTO DE IMAGENS</u> .....	3
2.1 - Amostragem e quantização .....	3
2.2 - Realçamento .....	3
2.3 - Restauração e reconstrução .....	6
2.4 - Codificação .....	9
2.5 - Análise de imagens .....	11
3. <u>CLASSIFICAÇÃO DE PADRÕES</u> .....	13
3.1 - Introdução .....	13
3.2 - Extração de atributos .....	14
3.3 - Métodos de classificação .....	15
4. <u>CONCLUSÕES</u> .....	18
BIBLIOGRAFIA .....	20

ABSTRACT

*This paper presents in a tutorial form a summary of the main areas of current activities in Image Processing and Pattern Classification.*

## 1. INTRODUÇÃO

Dentre as aplicações que a tecnologia digital possibilita no presente, destacam-se aquelas voltadas para a análise de dados bidimensionais, colhidos por diversos tipos de sensores. Tais aplicações têm recebido usualmente o nome de Processamento de Imagens. Este termo é reservado para as aplicações onde a entrada é uma imagem e a saída também é uma imagem. No caso de reconhecimento de padrões, a entrada do processo é uma imagem, mas a saída é uma classificação ou descrição da mesma, embora, freqüentemente, uma tarefa do tipo processamento de imagens (pré-processamento, em geral) venha a ser executada como parte do reconhecimento. Já no caso das chamadas aplicações gráficas por computador ("computer graphics"), a entrada do processo é uma descrição da imagem, a qual é obtida como saída.

O processamento de imagens tem tido um papel importante em áreas como: detecção de recursos naturais por meio de fotografias aereospaciais; análise de imagens meteorológicas; transmissão de imagens de sondas interplanetárias e exploração das possibilidades de televisão digital e de fac-símile; análise de imagens biomédicas, como em tomografia computadorizada ou de células sanguíneas e cromossomas; detecção de falhas em peças por meio de radiografias industriais, etc.

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 aereospaciais, 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 áreas de Processamento de Imagens e Reconhecimento de Padrões têm apresentado enorme progresso nos últimos vinte anos, e algumas dezenas de livros têm sido publicados sobre esses assuntos. Além disso, vários volumes têm sido editados, organizando os inúmeros artigos apresentados em conferências e publicados em revistas. Dessas últimas há três delas dedicadas especialmente a estas áreas. Para uma introdução, o leitor poderá consultar as referências mencionadas na bibliografia, no final deste trabalho.

As técnicas de processamento de imagens derivam de um grande número de disciplinas, incluindo, entre outras, Ótica, Neurofisiologia, Colorimetria, Engenharia Elétrica e Ciência da Computação. Este artigo se concentra nos aspectos ligados às duas últimas disciplinas, orientando a exposição para uma descrição sumária das técnicas envolvidas, e evitando maior detalhamento matemático. O mesmo pode ser dito sobre a exposição de classificação de padrões, onde não são cobertos os aspectos específicos da aplicação em cada área, mas as técnicas de uso geral. Também não são abordados os sistemas de "hardware" para processamento de imagens e reconhecimento de padrões, como "displays" e rastreadores, e nem as técnicas de processamento ótico, que também podem ser usadas nessas disciplinas.

## 2. PROCESSAMENTO DE IMAGENS

### 2.1 - AMOSTRAGEM E QUANTIZAÇÃO

Para processar uma imagem por computador, a função de duas variáveis que caracteriza o brilho é normalmente amostrada numa grade regular, e os valores resultantes são discretizados num número finito de níveis de cinza. Imagens a cores requerem três componentes (azul, verde e vermelho, por ex.), que são amostradas e quantizadas de maneira análoga. Tipicamente, uma imagem consiste numa matriz de 512x512 elementos denominados "pixels" (abreviatura do inglês "picture elements"), cada um possuindo 64 níveis de cinza.

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, apresentando transições bruscas de tonalidade, que originalmente eram suaves.

### 2.2 - REALÇAMENTO

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

É freqüente ocorrer a situação em que, após a digitalização da imagem (amostragem e quantização), os níveis de cinza ocupam 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 (utilizando saturação, por ex.) 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 uniforme ou ainda exponencial ou hiperbólica, levando em conta as características do sistema visual humano.

Uma imagem é freqüentemente 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 aplicadas simplesmente substituindo-se o ponto central de uma vizinhança  $3 \times 3$  ou  $5 \times 5$  de uma imagem pela média nessa vizinhança ou então envolvendo-se a imagem original com máscaras de coeficientes que possuam a característica de atenuar as altas freqüências, onde o ruído é geralmente preponderante sobre o sinal. 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 detalhes. Uma das operações mais simples para conseguir tal efeito é sub

trair 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 fotógrafos 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 operadores 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 o brilho da imagem é encarado 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 (intersecções de estradas, em imagens aéreas, por exemplo) 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 brilho 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-se técnicas de registro de imagens, a serem abordadas na seção de análise de imagens.

### 2.3 - RESTAURAÇÃO E RECONSTRUÇÃO

Enquanto que 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 imagem 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 ou imagens de satélites militares de alta resolução) 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-se as condições ideais de linearidade e invariança 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 aproximá-lo;

- 2) A inevitável presença do ruído faz com que o resultado possa ficar completamente mascarado, nas frequências espaciais nas quais o ruído (geralmente contendo 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 inerentemente 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.

Esse conhecimento pode ser de dois tipos: 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 2.4 sobre codificação.

A possibilidade de usar conhecimento determinístico "a priori" para combater o mal condicionamento também pode ser explorada

na forma de restrições de desigualdade. Uma restrição natural que ocorre em imagens reside no fato de valores de brilho 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-se a Transformada Rápida de Fourier. Tem havido recentemente tentativas de implementar métodos de restauração não-linear, levando-se 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 do problema 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  $\theta$ , o que foi obtido somando-se 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 pla

no 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.

#### 2.4 - CODIFICAÇÃO

O objetivo 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, causando 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-se 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, amostrando-se espaçadamente as componentes da Transformada de Fourier da imagem em baixas frequências, o mesmo sendo feito em 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), ou 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.

Nesse tipo de codificação, o fato de a transformação da imagem ser compressível é fundamental. Por exemplo, na Transformada de

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 distribuir os 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 do desempenho 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 que, simultaneamente, seja tratável do ponto de vista matemático e leve em conta as características complexas do sistema visual humano.

## 2.5 - ANÁLISE DE IMAGENS

A área de análise de imagens engloba várias subáreas mais ou menos independentes, que serão tratadas brevemente nesta seção.

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 filtração 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.

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 aeroespaciais de sensoriamento remoto ou meteorologia até imagens biomédicas. A primeira tentativa para resolver este problema consistiu em se efetuar uma correlação entre as duas imagens, localizando-se 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-determinado, 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 seção, seria conveniente citar brevemente alguns outros problemas em análise de imagens que têm sido objeto de investigação: decomposição de uma imagem em suas componentes conectadas, percorrimto de bordas, código de cadeia, métodos de propagação e contração, e representação de regiões por esqueletos.

### 3. CLASSIFICAÇÃO DE PADRÕES

#### 3.1 - INTRODUÇÃO

As motivações para o desenvolvimento da área de classifi  
cação de padrões podem ser encontradas:

- 1) No desafio de reproduzir com o computador as habilidades de to  
mar 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 che  
ques 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 se  
res muito primitivos. Todavia, mesmo diante dessa insuficiência, tem ha  
vido êxito na resolução de diversos problemas de reconhecimento de pa  
drões, particularmente nas tarefas mais simples de classificação, ou se  
ja, a de atribuir um objeto a uma de diversas classes. Os métodos desen  
volvidos têm tido por base, em geral, uma formulação estatística que po  
de estar distante dos processos perceptuais em seres vivos.

Deste modo, é feito, inicialmente, um conjunto de medidas por um sensor. Não serão discutidos neste artigo os problemas referen  
tes às medidas que devem ser tomadas nem ã maneira como deve ser proje  
tado 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 necessã  
rio efetuar um processo de redução da dimensionalidade (seleção de atri  
butos) antes da fase de classificação, de modo a reduzir o esforço com  
putacional 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 caem sempre na partição correspondente. Como esse objetivo ideal nem sempre pode ser alcançado, deve-se estabelecer a partição minimizando-se 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 aprendizado estatístico, uma vez que nas situações reais é raro ter uma descrição estatística completa das densidades de probabilidade que caracterizam cada classe.

### 3.2 - 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 de 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.

### 3.3 - 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 custos de decisão são escolhidos, o problema se enquadra na chamada teoria de decisão bayesiana, em quanto que se as classes são assumidas 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 existente 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 que 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 funções discriminantes, sem passar pelas densidades condicionais. Esses procedimen

tos 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 as sim permanecerão até o fim do mesmo. Os exemplos mais conhecidos de agregação hierárquica ocorrem em Biologia, onde indivíduos se agrupam em es pécies, espécies em gêneros, gêneros em famílias, etc. Apesar 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, de senvolve-se atualmente um esforço considerável no sentido de validar es taticamente 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 tare fa 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  $N$  passa gens, em cada uma das quais uma observação é mantida para teste, e as restantes  $N-1$  são usadas para treinamento.

#### 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

- 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*. New York, N.Y. Academic, 1976.

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

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