



PALAVRAS CHAVES/KEY WORDS		AUTORIZADA POR/AUTHORIZED BY	
AUTORES AUTHORS	PROCESSAMENTO DIGITAL DE IMAGENS	Dr. <i>Luis Gylvan Meira Filho</i> Diretor OBT	
AUTOR RESPONSÁVEL RESPONSIBLE AUTHOR		REVISADA POR / REVISED BY	
<i>Nelson Mascarenhas</i> Nelson D.A. Mascarenhas			
DISTRIBUIÇÃO/DISTRIBUTION		DATA / DATE	
<input type="checkbox"/> INTERNA / INTERNAL <input checked="" type="checkbox"/> EXTERNA / EXTERNAL <input type="checkbox"/> RESTRITA / RESTRICTED		Dezembro de 1990	
CDU/UDC		ORIGEM ORIGIN	
621.376.5		DPI	
PUBLICAÇÃO Nº PUBLICATION NO		PROJETO PROJECT	
INPE-5212-MD/46		TASIM	
TÍTULO/TITLE	INTRODUÇÃO AO PROCESSAMENTO DIGITAL DE IMAGENS	Nº DE PAG. NO OF PAGES	ULTIMA PAG. LAST PAGE
AUTORES/AUTHORSHIP	Nelson Delfino d'Ávila Mascarenhas	35	34
VERSÃO VERSION		Nº DE MAPAS NO OF MAPS	
VERSÃO VERSION		Nº DE MAPAS NO OF MAPS	
RESUMO - NOTAS / ABSTRACT - NOTES			
<p>As seguintes técnicas, que são utilizadas em processamento digital de imagens, são brevemente cobertas neste artigo: 1) amostragem e quantização; 2) realce; 3) restauração; 4) reconstrução; 5) codificação; 6) análise de imagens; 7) classificação de imagens.</p>			
OBSERVAÇÕES / REMARKS			
<p>Trabalho preparado como notas de aula de mini-curso oferecido na Jornada EPUSP/IEEE em Computação Visual, em São Paulo, SP, no período de 04 a 07 de dezembro de 1990.</p>			

INTRODUÇÃO AO PROCESSAMENTO DIGITAL DE IMAGENS

Nelson Delfino d'Ávila Mascarenhas
Instituto Nacional de Pesquisas Espaciais-INPE
Caixa Postal 515 - 12201 - São José dos Campos, SP
E-mail: INPEDPI@BRFAPESP.BITNET

ABSTRACT - The following techniques, which are used in digital image processing, are briefly presented in this paper: 1) sampling and quantization, 2) enhancement, 3) restoration, 4) reconstruction, 5) coding, 6) image analysis; 7) image classification.

1. INTRODUÇÃO

A tecnologia digital possibilitou, nos últimos vinte e cinco anos, um grande desenvolvimento das técnicas voltadas para a análise de dados multidimensionais, adquiridos por diversos tipos de sensores. Tais aplicações têm recebido usualmente o nome de Processamento Digital de Imagens. Este termo é reservado para as situações onde a entrada do processo é uma imagem e a saída é também uma imagem. No caso de reconhecimento de padrões, a entrada pode ser uma imagem, mas a saída é uma classificação ou descrição da mesma, embora, usualmente, uma tarefa do tipo processamento de imagens (pré-processamento, em geral) venha a ser executada como parte do reconhecimento. Por outro lado, no caso das chamadas aplicações gráficas por computador ("computer graphics"), a entrada do processo é uma descrição da imagem, que é então obtida como saída do processo.

As aplicações de processamento digital de imagens abrangem uma ampla gama de domínios, incluindo, por exemplo, análise de imagens aeroespaciais para monitoramento de recursos naturais em sensoriamento remoto ou meteorologia; transmissão de imagens de TV, fac-símile, tele-conferência e videofone por via digital; análise de imagens biomédicas por radiografia digital, tomografia, medicina nuclear, ultrassom, microscopia de células ou cromossomos; automação industrial, envolvendo, por exemplo, visão robótica, radiografias industriais e testes não-destrutivos de materiais.

As técnicas de Processamento Digital de Imagens resultam de inúmeras disciplinas como Ótica, Neurofisiologia, Colorimetria, Engenharia Elétrica e Ciência da Computação. Este artigo tutorial abordará apenas os aspectos ligados às duas últimas áreas, e apresentará uma descrição sumária e

introdutória das técnicas computacionais, com pequeno detalhamento matemático. O leitor poderá encontrar maiores detalhes e resultados recentes na bibliografia citada no final deste artigo. 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 ou híbrido que não serão cobertas por este tutorial.

2. AMOSTRAGEM E QUANTIZAÇÃO

Uma imagem de um objeto real é, em princípio, contínua tanto na variação espacial como nos níveis de cinza. Para que uma representação dessa imagem possa ser feita num computador digital, é necessário discretizar a imagem tanto no espaço como na amplitude.

A discretização no espaço faz-se pelo processo de amostragem. A ferramenta fundamental para determinar a mínima razão de amostragem é o teorema de Shannon que na sua forma unidimensional pode ser expresso da seguinte maneira: se a Transformada de Fourier de uma função $f(t)$ anula-se para $|W| \geq 2\pi f_c$, então $f(t)$ pode ser reconstruída exatamente por amostras com uma separação de $\frac{1}{2f_c}$. Essa separação mínima é também chamada razão de Nyquist.

Se a amostragem não é suficientemente fina ocorrerá o chamado fenômeno de "aliasing", que no caso de sinais bidimensionais de imagens provoca o surgimento de estruturas espúrias na imagem, denominadas padrões de Moiré.

Através do processo de quantização, uma imagem com tons contínuos é convertida numa com tons discretos. Para armazenamento digital, cada tom é representado por um código binário. Os métodos de quantização procuram, em geral, otimizar um critério de erro entre a imagem original e a imagem quantizada. Normalmente, o critério escolhido é o erro médio quadrático. Para tanto, a função densidade de probabilidade dos tons de cinza é suposta conhecida. É possível então efetuar a derivação dos níveis de decisão (na imagem contínua) e de quantização (na imagem discreta). O resultado indica que os níveis de decisão devem ser escolhidos como o ponto médio dos níveis de quantização e que, por sua vez, os níveis de quantização são dados pelo valor esperado dos tons contínuos condicionado ao correspondente intervalo de decisão. No caso da distribuição uniforme os níveis de quantização também são o ponto médio dos níveis de decisão. Quando a densidade não é uniforme, os resultados indicam que a minimização do erro médio quadrático é obtida pela

colocação dos níveis de decisão (e conseqüentemente os de quantização) mais próximos nas regiões de altos valores de densidade de probabilidade. Esse resultado é conhecido na literatura como "tapered quantization". A representação de uma imagem contínua por uma imagem quantizada com um número pequeno de tons de cinza implica o aparecimento do fenômeno visual denominado falsos contornos, pelo qual em regiões da imagem contínua que eram suaves aparecem transições visíveis em tom de cinza.

Uma digitalização típica de uma imagem em tom de cinza consiste de 512 x 512 ou 1024 x 1024 elementos denominados "pixels", cada um quantizado com 6 bits (64 níveis) a 10 bits (1024 níveis). Imagens a cores requerem três componentes (por exemplo: azul, verde e vermelho) que são amostradas e quantizadas de maneira análoga. Também é usual trabalhar com imagens binárias, como em aplicações industriais ou em fac-símile de documentos.

3. REALCE

De modo geral, as técnicas de realce visam melhorar a qualidade da imagem sob os critérios subjetivos de avaliação de observador humano. A utilização desses critérios torna geralmente difícil uma formulação matemática precisa dos processos de realce que, como consequência, têm, muitas vezes, um caráter ad hoc.

Frequentemente, após a digitalização da imagem os níveis de cinza estão concentrados numa pequena parte da variação tonal disponível. Por um processo de modificação da escala de cinza consegue-se espalhar a distribuição desses tons, tornando a informação mais visível ao olho humano. Tais métodos, denominados manipulação de contraste, envolvem o mapeamento linear ou não linear (por exemplo utilizando saturação) dos níveis originais, através do uso de tabelas. Uma expansão da escala tonal poderá provocar a existência de níveis de cinza não ocupados na imagem digitalizada de saída, daí resultando o fenômeno de falsos contornos. Isso pode ser evitado pela utilização de um número maior de níveis de quantização na imagem de saída.

Uma outra técnica de alteração da distribuição dos tons de cinza permite uma formulação matemática mais precisa. Trata-se do problema de equalização do histograma, que pode ser especificado do seguinte modo: dada uma variável aleatória representando os tons de cinza da imagem, achar uma função dessa variável aleatória tal que a variável aleatória de saída seja uniformemente distribuída. Mostra-se que, dentre a classe das funções monotonicamente

crecentes, a função resultante é a função de distribuição da variável aleatória original ($F_X(x)$). Se uma função monotonicamente decrescente fosse escolhida, o resultado seria $(1 - F_X(x))$. Para ocupar a gama de tons de cinza disponíveis (por exemplo: 0 a 255), basta multiplicar a variável aleatória resultante por um fator apropriado. Pode-se também especificar uma distribuição $F_Z(z)$ de saída e nesse caso demonstra-se que o mapeamento é dado por $Z = F_Z^{-1}(F_X(x))$ onde $F_Z^{-1}(\cdot)$ é a função inversa. Funções de distribuição de saída comumente especificadas são a gaussiana, a exponencial ou do tipo hiperbólico, levando em conta as características do sistema visual humano.

Uma imagem é frequentemente afetada por ruído de 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, conduzindo ao projeto de filtros de Wiener ou de Kalman (filtragem recursiva) em duas dimensões. Recentemente tem havido grande interesse em modelar a imagem original por campos aleatórios markovianos, caracterizados por distribuições de Gibbs, conduzindo a processos de estimação por "simulated annealing". Entretanto, tais caracterizações são, muitas 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 esse problema por um tipo de filtragem não-linear denominado filtro da mediana em que, numa janela com um número ímpar de "pixels", o ponto central é substituído pela mediana dos valores observados.

As técnicas de realce de bordas têm como objetivo aumentar a visibilidade de contornos de baixo contraste e a percepção de detalhes. Se uma imagem é rastreada eletronicamente, o sinal resultante pode ser passado por um filtro passa-altas. Outra possibilidade para obter o efeito de realce de bordas é usar a técnica denominada "unsharp masking", que consiste essencialmente em subtrair da imagem uma versão borrada dela mesma. A idéia subjacente é que essa versão contém basicamente as componentes de baixa frequência da imagem, do que resulta uma filtragem passa-altas. Outras possíveis operações de realce de bordas consistem na aplicação direta de

filtragens passa-altas no domínio do espaço (por máscaras convolucionais) ou no domínio de frequência (utilizando muitas vezes a Fast Fourier Transform-FFT). É possível obter aproximações digitais dos operadores gradiente ou laplaciano que são utilizadas também na detecção de bordas a ser analisada com mais detalhes na seção sobre análise de imagens. A desvantagem que resulta em geral na aplicação desses operadores é um aumento relativo de ruído na imagem.

Através de modelo em que o processo de formação da imagem é dado pelo produto de um fator de iluminação e outro de refletância, pode-se realçar este último pelas chamadas técnicas de filtragem homomórfica. Em geral, a componente de iluminação caracteriza-se por variações espaciais lentas, enquanto a componente de refletância tende a variar mais abruptamente. Denotando o processo de iluminação por $i(x,y)$ e o de refletância por $r(x,y)$, a imagem será caracterizada por $f(x,y) = i(x,y) \cdot r(x,y)$. Tomando o logaritmo, obtém-se $\log f(x,y) = \log i(x,y) + \log r(x,y)$. Obtendo a Transformada de Fourier serão obtidos dois sinais aditivos que, pelas considerações anteriores podem ser separados pelos menos parcialmente no domínio da frequência. A operação de Transformada de Fourier inversa seguida do processo de exponenciação permitirá recuperar a iluminação e a refletância filtradas.

Uma técnica de realce de imagens utilizada em muitas aplicações é a de realce a pseudo-cor. O processo consiste em mapear tons de cinza em cores. A base do realce está na habilidade muito maior do olho humano em distinguir cores do que tons de cinza. Contudo, a escolha do mapeamento tem sido, em geral, empírica. Têm contribuído para isso as dificuldades de modelagem matemática da percepção visual em cores do ser humano.

No caso de imagens multiespectrais frequentemente utilizadas em sensoriamento remoto, algumas operações de realce têm sido utilizadas. Dentre elas se destacam: 1) a geração de imagens que envolvem a razão entre bandas ou entre diferenças ou somas de bandas dessas imagens, com aplicações principalmente em geologia; 2) a transformação de principais componentes (também denominada Karhunen-Loeve ou Hotelling) que consiste em efetuar uma combinação linear das bandas através da multiplicação do vetor que caracteriza os "pixels" da imagem multiespectral em cada ponto por uma matriz cujas linhas são os vetores próprios da matriz de covariância espectral. Demonstra-se que as componentes resultantes, além de serem não correlacionados, podem ser ordenadas na ordem decrescente dos valores próprios correspondentes. As três principais componentes são então atribuídas às cores primárias verde, azul e vermelho; 3) a transformação intensidade - matiz - saturação (em inglês IHS:

intensity-hue-saturation), através do qual três bandas de uma imagem multispectral podem ser associadas às cores falsas vermelho (R), verde (G) e azul (B) e, a partir daí, três novas bandas são geradas pela conversão RGB - IHS.

4. RESTAURAÇÃO

O problema de restauração consiste na tarefa de estimar uma imagem que sofreu um processo de degradação, envolvendo algum tipo de espalhamento da radiação e/ou contaminação por ruído. Tais problemas advêm ou da resolução limitada dos instrumentos ou em condições adversas, como movimento relativo do sensor em relação ao objeto ou turbulência do meio de propagação da radiação. As aplicações abrangem um grande número de áreas variando de Astronomia Ótica, Radioastronomia, Aplicações Espaciais, Medicina e Biologia, etc.

Alumas simplificações são normalmente feitas, a fim de que o problema torne-se matematicamente tratável. A primeira delas consiste em admitir que o processo de degradação possa ser representado por um sistema linear seguido de uma adição de ruído independente do sinal. Embora simplificado, tal modelo tem se revelado satisfatório do ponto de vista experimental em muitas situações. Desta maneira, o espalhamento da radiação é caracterizado por uma função de espalhamento $h(.,.,.,.)$ de tal modo que o modelo de degradação é expresso pela Equação Integral de Fredholm do 1o. tipo que tem a forma

$$g(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha,\beta) h(x,\alpha,y,\beta) d\alpha d\beta .$$

No caso particular em que a função de espalhamento pode ser considerada invariante no espaço, resulta que $h(x-\alpha,y-\beta)$, e o processo de degradação pode ser representado por uma integral de convolução.

Em algumas situações é possível obter um modelo físico para o processo de degradação. Por exemplo, no caso de movimento uniforme unidirecional, a função de espalhamento tem a forma retangular e sua Transformada de Fourier é do tipo $H(u,v) = T \text{ sinc}(uVT)$ onde V é a velocidade e T é o tempo de exposição. Em outras situações, entretanto, um modelo físico para o espalhamento não é disponível e é necessário recorrer então a medidas na própria imagem degradada. Nesse caso o conhecimento a priori da existência de fontes aproximadamente pontuais, retas ou bordas retilíneas na imagem pode ser utilizado para se determinar a função de espalhamento.

Nas condições de linearidade e invariância do espaço, é válida a equação de convolução para modelar o espalhamento da radiação. Assim sendo, tomando 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. Na ausência de ruído haveria portanto uma indeterminação.

2) A inevitável presença do ruído faz com que o resultado tenda a ficar mascarado nas frequências espaciais onde 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 malcondicionado, 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, em geral, 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 rugosidade da solução, envolvendo, por exemplo, derivadas de segunda ordem 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 torna-se mais difícil pela maior 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 no domínio da frequência pelo uso da

Transformada Rápida de Fourier ("FFT-Fast Fourier Transform") ou aproximação no domínio do espaço, por deconvolução local da imagem.

A possibilidade de usar conhecimento determinístico a priori para combater o malcondicionamento também pode ser explorada na forma de restrições de desigualdades. 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.

Outra possibilidade consistiria no uso de conhecimento a priori estatístico sobre a imagem, recaindo-se portanto num esquema bayesiano de estimação. Se apenas operações lineares são consideradas, ou se sinal e ruído têm distribuição gaussiana e qualquer operação é permitida, o procedimento ótimo é o filtro de Wiener, que também pode ser implementado de maneira eficiente usando-se a Transformada Rápida de Fourier.

Um grande número de técnicas de restauração tem sido proposto recentemente. Pode-se mencionar, por exemplo, as técnicas não-lineares, para lidar com as não-linearidades associadas ao processo de gravação de imagens ou à presença de ruído dependente de sinal, como granulação fotográfica, ruído Poisson (característico de imagens com baixa intensidade de radiação) ou ruído "speckle", que ocorre em imagens obtidas por radiação coerente como laser ou radar de abertura sintética. Nesses casos, utilizam-se frequentemente técnicas adaptativas ou locais, envolvendo algumas vezes transformações que tornam o ruído independente do sinal. A técnica de restauração por máxima entropia baseia-se na analogia entre uma função densidade de probabilidade e a função $f(x,y)$ que representa a distribuição dos tons de cinza de uma imagem. Em ambos os casos as funções são não-negativas e o método garante igualmente que o resultado da restauração seja não-negativo.

5. RECONSTRUÇÃO

O problema de reconstruir uma imagem por suas projeções tem sido objeto de intensas investigações nos últimos anos, em particular devido à sua aplicação na área de tomografia computadorizada.

Deve-se notar que, do ponto de vista matemático, os problemas de restauração e reconstrução têm uma estrutura muito semelhante. Num modelo bem

simplificado, em ambos os casos, as observações resultam de uma combinação linear dos valores dos "pixels" da imagem original, além da inevitável presença do ruído.

O problema essencial em reconstrução de imagens é a recuperação da informação contida numa função de duas variáveis $f(x,y)$ por projeções dessa função em vários ângulos. Assim, em tomografia computadorizada por transmissão, o paciente é submetido à dosagem de raios-X, por exemplo, segundo vários ângulos que atravessam uma fatia de seu corpo. A partir dos dados obtidos, é possível reconstruir a densidade de tecido naquela fatia.

Um grande número de técnicas têm sido propostas para abordar o problema de reconstrução de imagens. De maneira geral, elas têm-se dividido em duas direções: a primeira delas utiliza as técnicas de transformadas da imagem e obtém seus métodos no domínio contínuo, discretizando os resultados no final para implementação digital; a segunda direção (constituindo as chamadas técnicas algébricas) parte inicialmente de um modelo já discretizado, reduzindo o problema à solução de um grande sistema de equações lineares.

Uma das técnicas do primeiro tipo é baseada no chamado Teorema do Corte de Fourier. Ele estabelece basicamente que a Transformada de Fourier de uma imagem bidimensional pode ser obtida pelas Transformadas de Fourier unidimensionais da imagem em várias direções, que são obtidas pelas projeções. Na prática, as Transformadas de Fourier contínuas em cada direção são discretizadas e calculadas pela Transformada Rápida de Fourier (FFT). Resta contudo o problema de que as projeções unidimensionais conduzem a amostras da Transformada de Fourier dispostas radialmente, ao passo que, para efetuar a Transformada de Fourier Inversa Discreta unidimensional para recuperar a imagem original, as amostras devem estar dispostas numa grade retangular, surgindo assim a necessidade de interpolação, que pode introduzir distorções, particularmente nas altas frequências da imagem.

A outra técnica do primeiro tipo é a mais utilizada nos tomógrafos comerciais modernos e consiste no algoritmo denominado convolução e retroprojeção (ou filtragem e retroprojeção). A base intuitiva deste método pode ser entendida tomando duas projeções em ângulos retos. Pode-se contruir uma função $f(x,y)$ que é constante numa dimensão e tem corte na dimensão ortogonal igual ao valor da projeção naquela direção. Se for feita uma soma das "retroprojeções" nos dois ângulos separados por 90 graus, a função bidimensional resultante será uma aproximação do objeto original. Para

melhorar a aproximação, cada projeção deve ser convolvida anteriormente com uma função apropriada. Esta convolução pode ser efetuada eficientemente no domínio da frequência através do uso da FFT.

As técnicas algébricas de reconstrução partem de um modelo já discretizado, em que a imagem bidimensional a ser reconstruída é representada por um conjunto de $N \times N$ "pixels". O valor observado no detector devido a uma projeção é o resultado da combinação linear dos valores dos "pixels" atravessados pela radiação onde os pesos podem representar o tamanho da intercessão do raio com a célula representada pelo "pixel". Obtém-se, assim, considerando o conjunto de todas as projeções, um sistema linear de equações, de grande dimensionalidade, mas ao mesmo tempo esparso, devido ao fato de, para um dado raio, apenas uma pequena percentagem dos "pixels" ser atravessada por ele. As técnicas algébricas consistem portanto em métodos de resolução desses grandes sistemas lineares esparsos. Tais métodos são geralmente iterativos, como o ART ("Algebraic Reconstruction Technique"). Estas técnicas, são usualmente bem mais lentas do que os métodos de convolução e reprojeção. Isso levou as tomógrafos comerciais de transmissão de raios X a utilizar quase que exclusivamente este último método para a reconstrução. Todavia, entre as vantagens apresentadas pelos métodos algébricos, pode-se citar a possibilidade de lidar com conhecimento a priori sobre a imagem, como a não-negatividade, por exemplo.

Os primeiros tomógrafos desenvolvidos utilizavam um feixe paralelo de raios para cada projeção. O conjunto de fonte emissora e sensores sofria uma rotação, novos dados eram coletados e assim sucessivamente. As máquinas mais modernas são baseadas em feixes divergentes, na forma de leque, para abreviar o tempo de exposição que pode ser crucial para reconstruir órgãos do corpo humano onde o problema de imobilidade é crítico. De modo geral, é possível estender as técnicas matemáticas desenvolvidas no caso de raios paralelos para a situação de raios divergentes.

Os principais métodos de reconstrução médica na atualidade são: a) tomografia de transmissão, utilizando raios X (utilizados também em testes não-destrutivos de materiais); b) ressonância magnética nuclear, com a vantagem de ser uma técnica não-invasiva; c) tomografia de emissão, cuja radiação provém da utilização de radioisótopos administrados ao paciente.

A tomografia da emissão propõe problemas de filtragem de ruído Poisson devido ao processo de contagem de partículas inerentes ao método. Tais

problemas podem ser resolvidos por critérios de otimalidade como, por exemplo, máxima verossimilhança, implementada pelo algoritmo EM ("Expectation-Maximization"). Pode-se, como alternativa, efetuar filtragens unidimensionais das projeções ruidosas e utilizar métodos convencionais de reconstrução (como filtragem-retroprojeção), abreviando bastante o esforço computacional.

6. CODIFICAÇÃO

A tendência para a utilização de tecnologias digitais em comunicações tem provocado um grande interesse na codificação digital da informação pictórica. Os exemplos abrangem um grande número de áreas, incluindo TV digital, sistemas de teleconferência, videotelefonia, fac-símile, etc. Além disso, abrem-se possibilidade de armazenamento e transmissão de imagens médicas, imagens e mapas de previsão do tempo, gráficos de engenharia, arquivos de impressões digitais, etc.

De um ponto de vista bem amplo, o objetivo dos processos de codificação de imagens é reduzir o volume de dados necessários para representar uma imagem com a finalidade de diminuir o tempo, a largura de faixa ou a potência necessárias para transmitir a imagem ou a quantidade de memória para armazená-la.

Os métodos de codificação de imagens baseiam-se tanto na redundância existente nas imagens quanto nas limitações do sistema visual humano. Embora ainda se esteja longe de um conhecimento completo sobre esse sistema, é possível explorar as limitações de resolução radiométrica, cromática, espacial ou temporal da visão humana.

As técnicas de codificação podem ser classificadas em dois tipos: reversível e irreversível, ou seja, no primeiro caso é possível uma exata reconstrução da imagem original, enquanto no segundo caso alguma distorção é tolerada. Entre as técnicas reversíveis ou exatas deve-se mencionar o código de Huffman que associa palavras mais curtas aos tons de cinza mais frequentes na imagem, minimizando assim a taxa de bits. Códigos estatísticos do tipo de Huffman têm sido aplicados em situações onde não se deve ter perda de informação, por exemplo, por razões legais, como em imagem médicas. Nas situações onde a taxa de compactação é maior deve-se tolerar algum tipo de distorção e os códigos estatísticos podem ser usados apenas como uma fase do processo de codificação.

A base matemática para a codificação de imagens é fornecida pela teoria de razão de distorção, iniciada por Shannon. Todavia, esta teoria fornece apenas os limites dos métodos de codificação não especificando como projetá-los. Além disso, ela sofre das deficiências na modelagem estatística realística da fonte de sinal (imagens), das dificuldades de cálculo da função da razão de distorção (mínima taxa de bits necessária para transmissão ou armazenamento sob uma dada distorção) e da necessidade de estabelecer um critério de fidelidade que, simultaneamente, seja tratável do ponto de vista matemático e leve em conta as características complexas do sistema visual humano.

Assim sendo, têm sido desenvolvidos métodos subótimos de compressão de imagens, que podem ser classificados em três tipos: a) codificação espacial; b) codificação por transformadas e c) codificação por modelos de imagens. Tais métodos serão brevemente descritos nesta seção.

Os métodos práticos de codificação digital envolvem em geral três estágios: 1) um estágio inicial no qual é feito algum tipo de transformação da imagem, em geral reversível, com o objetivo de reduzir a redundância estatística; 2) um estágio, usualmente irreversível, em que a acurácia da representação é reduzida mantendo-se uma qualidade visual da imagem; 3) um estágio reversível no qual a codificação propriamente dita é realizada, podendo ser utilizados os códigos estatísticos como os de Huffman.

Dentre os códigos espaciais, o mais simples é a Modulação com Codificação por Pulsos ("Pulse Code Modulation"-PCM). Neste sistema o sinal é amostrado (em geral na taxa de Nyquist) e quantizado num número de tons de cinza (M) que usualmente é uma potência de 2 ($M=2^b$) resultando numa codificação em b bits. O PCM não elimina qualquer redundância nos "pixels" da imagem e pode ser considerado um sistema de referência, em relação ao qual os demais sistemas são comparados em termos de taxa de transmissão.

A redução no número de níveis de quantização de uma imagem codificada por PCM leva ao efeito já mencionado de falsos contornos, que surgem em regiões de transição suave da iluminância. Esse efeito pode ser reduzido substancialmente pela adição de um ruído pseudo-aleatório antes da quantização e subtraindo-o sincronamente no receptor.

Os códigos preditivos têm suas raízes na teoria de séries temporais desenvolvida por Wiener na década de 40. A idéia básica é codificar o sinal

resultante da diferença entre o "pixel" e sua predição, baseada na mesma linha (predição unidimensional), em linhas adjacentes (predição bidimensional) ou mesmo envolvendo "pixels" nos quadros precedentes, no caso de imagens de TV. A alta correlação normalmente existente entre "pixels" adjacentes de uma imagem faz com que o histograma do resíduo de predição seja concentrado ao redor do valor zero. A densidade de Laplace tem sido utilizada para modelar esse histograma e um quantizador escalar ótimo pode ser utilizado, resultando em pequeno espaçamento para pequenos erros de predição e grande espaçamento para grandes erros. A codificação poderia ser feita utilizando códigos de comprimento variável como os de Huffman, embora, em geral, sejam empregados códigos com palavras de comprimento uniforme.

Ao sistema de codificação por predição dá-se, de modo geral, a denominação Modulação por Codificação Diferencial de Pulsos ("Differential Pulse Code Modulation"-DPCM). No caso particular de 2 níveis, o sistema de codificação tem o nome de Modulação Delta ("Delta Modulation"-DM). Um sistema DPCM com quantizador ótimo de 8 níveis (3 bits) pode produzir imagens com a mesma qualidade visual que um sistema PCM com 6 a 8 bits, com exceção dos erros em bordas agudas da imagem.

Os dois principais problemas nas imagens codificadas com DPCM são: a) sobrecarga, que ocorre em bordas cuja variação do sinal é mais alta que o intervalo de quantização; b) ruído de granulação, que resulta de uma oscilação entre dois níveis adjacentes de quantização em regiões uniformes da imagem.

Códigos preditivos bidimensionais oferecem em geral uma pequena redução no erro médio quadrático em comparação com os códigos unidimensionais. Todavia, com o uso de predição bidimensional, as bordas verticais tendem a ser mais bem representadas, bem como os efeitos de erros de transmissão.

Imagens tendem a exibir uma forte não estacionariedade estatística, o que tem motivado o uso de códigos adaptativos. Várias possibilidades têm sido exploradas, entre as quais a adaptação dos níveis de quantização dos códigos preditivos ao conteúdo da imagem. Se o brilho varia rapidamente, os níveis de quantização devem ser aumentados para reduzir a sobrecarga e, se o brilho é relativamente constante, os níveis de quantização devem ser diminuídos para combater os efeitos de ruído de granulação.

A redundância existente entre quadros sucessivos de TV pode ser explorada por técnicas preditivas. Todavia, devido à dificuldade de armazenamento, tais códigos, em geral, têm sido baseados na predição apenas pelo quadro anterior.

Um dos problemas que surgem com a utilização de códigos preditivos é o efeito de erros de transmissão. No receptor um tipo de integração é feito sobre os sinais de resíduo, de modo que um erro de transmissão tem a tendência de se propagar ao longo das linhas nas quais é feita a predição. Se o preditor é estável, os efeitos dos erros de transmissão tendem a decair. Pode-se aliviar esses efeitos reinicializando-se periodicamente o preditor pela transmissão do "pixel" original ou pela aplicação de um filtro de mediana unidimensional ao longo das colunas no receptor.

Um outro tipo de codificador espacial é o chamado codificador em dois canais, no qual a imagem é dividida em duas componentes, a de baixas frequências e a de altas frequências. A primeira componente pode ser bastante sub-amostrada e a segunda componente pode ser quantizada com grande espaçamento. O codificador de dois canais é um caso especial da chamada codificação por sub-bandas, onde cada componente a ser codificada é a saída de um filtro passa-faixas sobre a imagem original. Tipicamente, têm sido utilizados 16 filtros passa faixas, com um codificador projetado para cada saída dos filtros.

Uma variedade de codificação espacial conceitualmente relacionada à anterior é a chamada codificação por pirâmide. Uma possível representação por pirâmide consiste na imagem original e sucessivas imagens borradas e de menor resolução derivadas da primeira. No caso mais simples no qual a pirâmide possui apenas um nível, a imagem original é convolvida com um filtro passa baixas (por exemplo do tipo gaussiano e separável), gerando uma imagem borrada que é amostrada em mais baixa resolução. A diferença entre a imagem original e a imagem borrada é quantizada com menor número de níveis por conter essencialmente informação de alta frequência. Essa estrutura pode ser refinada gerando-se sucessivas imagens de menor resolução e as correspondentes imagens de erro de predição. No receptor percorre-se o caminho inverso, isto é, reconstrói-se a imagem original a partir daquela de menor resolução e dos sinais de erro de predição correspondentes, que foram transmitidos. Uma das vantagens desse método é a possibilidade de transmissão progressiva da imagem, interrompendo-se o processo quando a resolução estiver satisfatória.

A codificação por transformada baseia-se na idéia de aplicar uma transformação linear sobre a imagem ou sobre uma subimagem (para diminuir a carga computacional), de modo a obter componentes transformadas menos correlacionadas do que os "pixels" originais. Em seguida, estas componentes são ordenadas segundo sua importância em representar a informação, e as componentes menos representativas são descartadas (ou quantizadas com menor número de bits) para obter compressão. No receptor a transformação inversa é feita para recuperar a imagem.

A Transformada de Karhunen-Loeve ou de Hotelling (já mencionada na seção sobre realce) possui a característica de ser ótima no sentido de decorrelacionar as amostras, bem como de minimizar o erro médio quadrático de reconstrução, quando se descartam os coeficientes menos representativos.

O uso da Transformada de Karhunen-Loeve apresenta vários problemas na prática: 1) como ela é uma transformada que depende das características estatísticas da imagem, que frequentemente são não-estacionárias, a transformada vai depender da localização do bloco a ser transformado na imagem; 2) a transformada depende da caracterização estatística da imagem, o que é frequentemente difícil; 3) em geral, essa transformada não possui algoritmos rápidos para sua implementação.

Assim sendo, várias outras transformadas subótimas têm sido utilizadas na compressão de imagens. Elas produzem coeficientes menos correlacionados que a imagem original, não dependem do conhecimento da estatística da imagem e podem ser implementadas por algoritmos eficientes. Dentre essas transformadas, pode-se mencionar a Transformada Discreta de Fourier, a Transformada Discreta de Cosseno, a de Hadamard, a de Haar e a Transformada Oblíqua ("slant").

A Transformada Discreta de Fourier pode ser implementada eficientemente através de algoritmos de FFT ("Fast Fourier Transform"). Como no caso das demais transformadas mencionadas, ela é separável, o que implica que a transformada pode ser obtida aplicando sucessivamente transformações unidimensionais sobre as linhas da imagem e, em seguida, sobre as colunas. Deste modo, a transformada bidimensional pode ser implementada com $2N \log_2 N$ multiplicações e adições, onde N é o número de "pixels" ao longo de uma linha (ou coluna).

A Transformada do Cosseno tem sido objeto de grande interesse devido à sua proximidade das propriedades ótimas da Transformada de Karhunen-Loeve, ao

fato de poder ser implementada por um algoritmo mais rápido do que a FFT, que envolve apenas operações reais, e à redução do efeito visual de blocos na implementação da codificação por subimagens.

Depois da transformada, as amostras são selecionadas para transmissão ou armazenamento segundo dois tipos de estratégias usuais:

a) amostragem por zona, quando apenas os coeficientes que se situam numa região do plano transformado, em geral associada às frequências mais baixas, são codificados. A degradação pode ser grande se a imagem contém grandes componentes que são descartadas. No receptor, os coeficientes eliminados são usualmente iguados a zero.

b) amostragem por limiar, quando um coeficiente é transmitido se sua magnitude ultrapassa um limiar especificado. Nesse caso é necessário transmitir, em adição aos coeficientes selecionados, a informação da posição desses coeficientes.

A quantização dos coeficientes transformados tem sido baseado em modelos estatísticos escalares, utilizando, por exemplo, o esquema com base no mínimo erro quadrático. Como os coeficientes tendem a ser bastante decorrelacionados, usualmente não têm sido utilizados os mais recentes quantizadores vetoriais, em que um bloco de variáveis é quantizado conjuntamente. Tais quantizadores exigem um esforço computacional relativamente grande no transmissor e têm sido mais utilizados em códigos espaciais para aplicações de videodifusão digital. O primeiro coeficiente é geralmente descrito pela densidade de Rayleigh e os demais pela densidade gaussiana. A atribuição do número de bits a cada coeficiente poder ser feita dividindo-se o número total de bits pelos coeficientes de modo que o erro médio quadrático seja minimizado.

O efeito dos erros de transmissão em codificação por transformadas é bastante diferente daquele que ocorre em códigos preditivos. Naquele método a imagem reconstruída é uma combinação linear dos coeficientes transformados, de modo que o efeito de erros em coeficientes distribuem-se por todo o bloco.

Uma comparação simplificada dos códigos espaciais e dos códigos por transformadas revela que os primeiros são de implementação mais simples e mais utilizados em taxas relativamente altas de razão bits/"pixel" (2 a 3 bits/"pixel") como exemplo em TV digital. Já os últimos são de implementação mais complexa mas tendem a ter desempenho melhor em taxas mais baixas (ao

redor de 0,25 a 1 bit/"pixel"), como em videotelefonia. Como um compromisso entre as vantagens e desvantagens de ambos os sistemas é possível utilizar a codificação híbrida predição-transformação, envolvendo por exemplo a transformada unidimensional ao longo das linhas da imagem, seguida de DPCM na direção vertical.

Recentemente tem havido um grande esforço de pesquisa na chamada codificação por modelos de imagens. Por este método, a imagem ou uma porção da mesma é modelada e os parâmetros do modelo são transmitidos e utilizados no receptor para a síntese da imagem. Um exemplo desse tipo de codificação é a representação de texturas por modelos autorregressivos bidimensionais. O potencial de redução de taxa de bits/"pixel" é muito grande mas o método exige, em geral, processos custosos de segmentação de imagens. As aplicações em vista são caracterizadas pelo fato de a inteligibilidade da imagem, mais do que a qualidade visual, ser o fator fundamental.

Imagens binárias de fac-símile podem ser codificadas por vários métodos, entre os quais os códigos de comprimento de corridas ("run length coding"). Através do rastreamento da imagem tipo TV tem-se um número relativamente pequeno de "corridas" com o mesmo tom e a imagem pode ser codificada pela especificação do tamanho dessas "corridas" e dos tons a elas associados.

Finalizando esta seção, é conveniente mencionar que vários tipos de codificação têm sido desenvolvidos para imagens com características particulares, como imagens multiespectrais ou em cores, ou ainda sucessivos quadros de imagens de TV, onde uma pequena fração de pontos sofre mudança de quadro para quadro.

7. ANÁLISE DE IMAGENS

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

Algoritmos para detetar bordas em imagens podem ser uma etapa de pós-processamento em vários sistemas de processamento de imagens e reconhecimento de padrões. As técnicas mais simples consistem num processo de realce de bordas linear ou não-linear, em geral consistindo em algum tipo de diferenciação local, seguido de uma comparação do resultado com um limiar. Existem vários problemas ligados a tais métodos que, em geral, têm uma formulação heurística. Pode-se citar a escolha do limiar, que, em geral, deve

ser variável com a região da imagem, e a acentuação do ruído inerente a um processo de diferenciação. Inúmeras máscaras convolucionais têm sido propostas como operadores lineares, algumas com características direcionais e outras, como as aproximações discretas do laplaciano, sem essas características. Os operadores não-lineares em geral estão limitados a processamento sobre janelas de 2×2 ou 3×3 ao redor do pixel considerado. Dentre os mais populares estão a operação de diferenciação cruzada de Roberts, o operador de Sobel e o de Kirsch. Frequentemente, os operadores não-lineares de detecção de bordas podem ser decompostos numa operação pontual não-linear, seguida de uma operação linear e da comparação com um limiar.

A sensibilidade ao ruído dos detetores de borda pode ser contornada de algumas maneiras. Dentre elas, pode-se citar o uso de filtros gaussianos passa-baixas em combinação o operador laplaciano (e a conseqüente localização das bordas pelos cruzamentos de zero) e a utilização da teoria estatística de detecção de sinais, pela formulação do problema em termos de testes de hipóteses. Alguns detetores morfológicos de borda, que envolvem operações de máximo e mínimo locais numa imagem, também têm se relevado relativamente imunes ao ruído.

Frequentemente os detetores de borda fornecem resultados que consistem em pontos isolados, não correspondendo à realidade física. É possível aliviar esse problema pela utilização do contexto espacial no processo de decisão. Isso pode ser feito utilizando várias versões dos chamados processo de relaxação. A extração de contornos fechados ou pelo menos contínuos pode envolver a aproximação de curvas se existe informação a priori a respeito do tipo de objeto. Na aproximação por segmentos de retas tem sido utilizada a Transformada de Hough, na qual os pontos de borda são transformados para outro espaço, com a propriedade de que os pontos alinhados se transformam em agregamentos. Este método pode também ser generalizado para curvas de maior grau, ao preço de um aumento de dimensionalidade no espaço transformado. Em geral, o problema de unir pontos de borda para obter contornos pode ser encarado como o de encontrar caminhos de custo mínimo num grafo. Os nós do grafo são dados pelos pontos de borda e os arcos entre eles têm um custo definido que é função da distância e da similaridade das propriedades das bordas, como a orientação, por exemplo.

A segmentação de uma imagem pode ser feita também por regiões, ao invés de contornos. A técnica mais simples envolve a limiarização dos tons de cinza. Os problemas que surgem podem advir do ruído presente na imagem, das

diferenças devido à iluminação não homogênea, ou do fato de objeto e fundo possuírem gama ampla de valores, no caso de binarização. Os procedimentos de limiarização envolvem muitas vezes experimentação e tentativa e erro. Alguns métodos analíticos de escolha do limiar têm sido propostos. Um deles consiste em escolher o limiar tal que o histograma acumulado até esse valor coincida com uma avaliação a priori da distribuição de probabilidade. Outro método estabelece o limiar como o ponto de mínimo de um histograma bimodal que reflete a presença de um objeto e do fundo em diferentes faixas dos tons de cinza. O problema que ocorre é o fato de o histograma ser geralmente bastante irregular. Desta maneira, o vale do histograma pode ser aproximado por polinômios, por exemplo, para determinar o mínimo. A variação dos tons de cinza que definem o objeto ou o fundo pode determinar a necessidade de desenvolver um esquema adaptativo para a escolha do limiar. No caso de imagens em cores ou multiespectrais o problema de limiarização torna-se multidimensional. Desta forma, técnicas de agregação de dados podem ser utilizadas para definir a superfície de decisão multidimensional que estabelece a segmentação. Em princípio, pela utilização da função discriminante linear de Fischer é possível projetar os dados multivariados numa reta e nela estabelecer o limiar.

Abordagens para segmentação de imagens baseadas diretamente na obtenção de regiões podem ser classificadas em dois tipos: a) método de crescimento de regiões. Nesta abordagem inicia-se o processo por pequenas regiões que possam ser consideradas homogêneas. Regiões vizinhas são agrupadas baseadas num critério de similaridade radiométrica ou outros como inclusão de uma região em outra ou se a região agrupada possui uma forma mais regular, por exemplo. O processo é repetido até que nenhuma nova região é formada. Para evitar que regiões bastante fragmentadas resultem pode-se usar conhecimento a priori sobre propriedades da imagem, como relações de adjacência entre regiões; b) método da divisão e fusão. Uma alternativa à abordagem anterior é iniciar o processo pela imagem inteira, subdividindo-a num conjunto de regiões disjuntas e, em seguida, dividir ou fundir regiões segundo diferentes critérios possíveis. Uma estrutura de dados frequentemente utilizada neste processo é a chamada árvore quaternária ("quadtree"), onde cada nó tem quatro descendentes, provenientes da subdivisão de um quadrado ou retângulo em quatro quadrantes.

A segmentação da imagem pode ser feita pela atribuição de um "pixel" ou grupos de "pixels" a uma determinada classe. Esse processo de classificação de imagens será descrito com mais detalhes na próxima seção.

Uma vez segmentada a imagem, pode-se descrever os contornos ou regiões por vários métodos. Um dos métodos frequentemente utilizados para descrição de contorno é o código de cadeia, no qual as direções dos vetores que ligam sucessivos "pixels" de contorno são codificadas, normalmente em 3 bits (oito direções), além do código da posição do "pixel" inicial. Dada uma figura no plano é possível definir uma função intrínseca da figura que especifica a curvatura do perímetro como função do perímetro medido a partir de um ponto arbitrário. Como a função intrínseca é periódica sobre o comprimento do perímetro, ela pode ser expandida em série de Fourier. Dada então uma figura, os primeiros coeficientes da expansão podem ser computados, constituindo uma descrição dessa figura. Se o perímetro não contiver descontinuidades abruptas de curvatura, os primeiros coeficientes da expansão de Fourier fornecerão uma boa representação de forma. Outra representação de contornos pode ser feita por splines, bastante utilizadas também em computação gráfica.

A representação de regiões pode ser feita de várias maneiras. Serão descritas brevemente a seguir três abordagens que têm sido utilizadas:

- propriedades métricas. Estes atributos são baseados na noção de uma medida de distância entre pontos no espaço bidimensional da imagem. Estabelecendo uma noção de distância, pode-se introduzir vários atributos de forma de uma região como perímetro (P), área (A) e, por exemplo, a razão de alongação, definida por $T = 4\pi A/P^2$, que assume o máximo valor 1 quando a região a ser descrita é um círculo;
- propriedades topológicas. Estas propriedades de forma são aquelas invariantes às chamadas distorções do tipo "lâmina de borracha". Duas das propriedades topológicas de um conjunto que são utilizadas na descrição de forma são o número de componentes conectadas (C) e o número de buracos (H), que são relacionados pelo chamado número de Euler (E) por $E = C - H$;
- propriedades analíticas. Uma possível descrição analítica de forma pode ser feita estabelecendo uma analogia entre uma função densidade de probabilidade bidimensional e a função $f(x,y)$ que representa a iluminação de uma imagem. Assim sendo, pode-se definir uma função característica pela Transformada de Fourier de $f(x,y)$ e utilizar as propriedades de geração de momentos daquela função. É de se esperar que a forma da região descrita por $f(x,y)$ possa ser representada pelos primeiros momentos dessa função.

A estrutura de objetos numa imagem pode caracterizá-los através de linhas ou arcos, por exemplo. Exemplos incluem caracteres impressos ou manuscritos, impressões digitais, cromossomos, etc. Em tais situações a largura das formas não contribui para sua descrição. Deste modo, têm sido desenvolvidos algoritmos para a obtenção do esqueleto de uma figura, bem como

algoritmos de afinamento de formas. As técnicas de morfologia matemática, muito úteis nessas situações, têm atraído considerável atenção recentemente.

Inúmeras técnicas de análise de textura em imagens têm sido desenvolvidas. O problema é relevante em imagens biomédicas ou de sensoriamento remoto, por exemplo. Uma definição precisa de textura ainda não foi proposta, embora seja poderosa a capacidade da visão humana em sua discriminação. Uma caracterização natural é feita através do espectro de Fourier, pois uma textura grossa poderia em princípio distinguir-se de outra mais fina pela maior concentração de energia da primeira em frequência mais baixas. A densidade espectral de potência também contém informações de bordas e retas na imagem. Todavia, os resultados experimentais obtidos até agora têm evidenciado que pode haver bastante sobreposição no domínio de Fourier para tipos diferentes de textura em imagens aéreas. Por outro lado, atributos obtidos do espectro de Fourier têm sido capazes de classificar imagens radiográficas pulmonares. A Transformada Inversa de Fourier da densidade espectral de potência é a função de correlação que, de maneira análoga, pode caracterizar a textura de uma dada região. Nesse caso, texturas mais grossas exibirão maior correlação para valores mais altos dos argumentos dessa função. A densidade de bordas locais por unidade de área também tem sido proposta como uma medida de rugosidade de uma textura. As chamadas matrizes de co-ocorrência têm sido bastante utilizadas. Elas consistem basicamente no histograma bidimensional de níveis de cinza para pares de "pixels" separados por uma dada distância e um dado ângulo. Uma textura fina pode ser considerada como um processo estocástico que tende a exibir baixa correlação entre "pixels" adjacentes e, como consequência, o histograma bidimensional tende a se espalhar, ao passo que uma textura mais grossa exibirá maior correlação entre "pixels" próximos, provocando uma concentração do histograma ao redor da diagonal principal. Os modelos de séries estatísticas uni e bidimensionais também têm sido utilizados tanto para a análise de textura como para sua síntese. A série unidimensional que resulta de um rastreamento tipo TV de uma seção de uma imagem exibindo uma dada textura possui características sazonais. Dentre os modelos numéricos de textura, devem ser mencionados ainda os modelos de mosaicos, que particionam a imagem em regiões de acordo com alguma lei probabilística e os mais recentes modelos baseados em fractais ou em campos aleatórios markovianos caracterizados pelas distribuições de Gibbs. Além das caracterizações numéricas de textura, é possível desenvolver também caracterizações sintáticas através de gramáticas de vários tipos.

A superposição ("registro") precisa de duas imagens obtidas em condições diversas pode ser necessária em aplicações que vão desde imagens aeroespaciais de sensoriamento remoto ou meteorologia até imagens biomédicas. Assim, o problema é importante a) quando se têm imagens obtidas por diferentes sensores e se deseja saber a resposta no mesmo ponto dos diversos sensores; b) quando se têm imagens obtidas em diferentes épocas ou instantes de tempo; c) quando se têm imagens tomadas de posições diferentes e se pretende obter informação tridimensional da cena; d) quando se deseja registrar uma imagem obtida por sensores com um mapa.

O primeiro passo a ser dado uma operação de registro quando há diferenças entre as geometrias das imagens é determinar a transformação geométrica que as relaciona. Se não houver informação adicional sobre a transformação a não ser as próprias imagens, isto é feito através da determinação de pares de pontos denominados "pontos de controle" nas duas imagens que correspondem ao mesmo objeto. Uma vez determinados esses pontos, é possível obter funções que mapeiam as coordenadas de uma imagem na outra. Normalmente são utilizados polinômios nas duas variáveis espaciais para esta finalidade.

O segundo passo consiste na construção da imagem registrada pela transformação geométrica. A transformação define as posições dos pontos e esquemas de interpolação definem o nível de cinza a ser atribuído a cada ponto. Essa interpolação normalmente é local e pode envolver por exemplo o "pixel" mais próximo apenas ("vizinho mais próximo"), os quatro "pixels" ao redor do ponto mapeado pela transformação geométrica, por uma função bilinear ou os 16 "pixels" vizinhos (convolução cúbica).

Muitas vezes, quando não há necessidade de rotação e as escalas são as mesmas, o registro reduz-se a achar uma translação conveniente. O procedimento normalmente usado é definir uma medida de similaridade, computá-la para cada translação e tomar a translação de maior medida. Essa é a filosofia que governa o método de correlação cruzada entre as duas imagens, localizando o ponto de máximo. As dificuldades residem na imprecisão da determinação do ponto ótimo devido a diferenças sazonais, ruído, etc, e na grande carga computacional, a despeito do uso do algoritmo de FFT para efetuar a correlação. Vários esquemas têm sido propostos para reduzir a carga computacional, guiados pela idéia de que um esforço maior deve ser dedicado apenas ao redor do verdadeiro ponto de registro. Um esquema que tem sido amplamente utilizado é o algoritmo de detecção sequencial de similaridade:

acumula-se o módulo do erro que resulta da comparação das duas imagens para uma posição candidata de registro. O acúmulo de erro é interrompido assim que este valor ultrapassa um dado limiar, que pode ser constante ou monotonicamente crescente. Pode ser derivada uma sequência de limiares segundo uma abordagem estatística baseada na distribuição exponencial do erro na posição de registro. A possibilidade de usar a ferramenta dos testes sequenciais de hipóteses também tem sido explorada tanto para imagens com tons de cinza como para imagens binárias, utilizando respectivamente modelo gaussiano e binomial. A idéia é que, no primeiro caso, o ponto de registro é caracterizado por uma baixa variância da diferença dos "pixels" das imagens e, no segundo caso, a probabilidade de o erro binário ser 1 também é baixa no ponto de registro.

8. CLASSIFICAÇÃO DE IMAGENS

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. 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 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 apenas as técnicas de classificação de padrões que têm sido mais utilizados em áreas como detecção de recursos naturais.

Os conhecimentos atuais de neurofisiologia não têm sido em geral suficientes para reproduzir, na máquina, as funções de percepção de seres muito primitivos. Todavia, mesmo diante desta 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 sido 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. No caso de imagens, estas medidas tendem a se adaptar ao modo de aquisição das imagens e ao tipo de classificação que se planeja desenvolver. Imagens de satélites de recursos naturais são colhidas em geral sobre várias bandas do espectro (tipicamente de 4 a 12) fornecendo assim um atributo natural para a classificação de um ponto no solo através do vetor de medidas nessas bandas. Outras possíveis medidas consistiriam em medidas de textura ou de tamanho e forma de objetos nessas imagens. Os dados obtidos são representados por meio de um vetor de medidas num espaço com uma determinada dimensão. Normalmente, é necessário efetuar um processo de redução de dimensionalidade (seleção de atributos) antes da fase de classificação, de modo a reduzir o esforço computacional nesta última fase.

A etapa de classificação reduz-se 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 apresentarem 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. Essas variações podem ser tanto devidas à incerteza inerente à caracterização da classe, como oriundas de perturbações estatísticas (ruído) nos sensores de medida.

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.

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. Estas técnicas podem ser divididas em: a) Transformação no Espaço de Atributos. O objetivo aqui é transformar o espaço de atributos 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 erro médio quadrático ou entropia. A Transformação de Karhunen-Loeve, 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, os quais 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 ser mais fáceis de computar do que outras medidas propostas.

O problema central de classificação, para um dado vetor de atributos observado, envolve basicamente o cômputo 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 enquadra-se na chamada teoria bayesiana, ao passo que se as classes são equiprováveis tem-se a decisão por máxima verossimilhança.

O modelo gaussiano tem sido utilizado frequentemente para caracterizar as classes, particularmente na área de sensoriamento remoto. Em geral, nestes casos, um limiar é também estabelecido para rejeitar padrões cuja distância ponderada - a distância de Mahalanobis - a cada uma das classes é superior ao limiar. Também têm sido utilizados os chamados classificadores de paralelepípedos em que cada classe é definida pelos limites inferior e superior do histograma obtido na área de treinamento. Tais classificadores podem ser implementados de maneira muito rápida por "hardware", mas seu desempenho tem deficiências, principalmente devido à ambiguidade que resulta da sobreposição de paralelepípedos relativos a duas ou mais classes.

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. Alguns resultados da aplicação desta técnica na classificação de caracteres alfanuméricos manuscritos tem sido utilizados.

Os métodos de classificação descritos a seguir referem-se essencialmente ao problema que ocorre quando as densidades de probabilidade que caracterizam as classes 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 conderavelmente 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, estas formas paramétricas têm a forma unimodal, enquanto os histogramas reais são muitas vezes multimodais. Deste modo, recorre-se às chamadas técnicas não-paramétricas. Algumas destas 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. Estes 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 pela necessidade de um número maior de amostras.

Suponha-se agora que sejam oferecidas ao classificador apenas amostras de classificação conhecida. Nestas condições o classificador deve ajustar iterativamente as superfícies de decisão de modo a, se possível, separar perfeitamente as amostras. Estes procedimentos, que tiveram muita popularidade no início da década de 60, são os algoritmos do tipo "perceptron" ou de relaxação. Essa linha de pesquisa foi retomada na década de 80, dando origem às chamadas redes neurais, que estão sendo intensamente investigadas atualmente. Para lidar com a situação em que as classes não são separáveis linealmente, 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. Este critério pode ser mínima variância, envolvendo as matrizes de espalhamento entre classes ou dentro de classes. Técnicas baseadas em teoria de grafos, na análise combinatória ou ainda na teoria de conjuntos nebulosos ("fuzzy sets") têm sido desenvolvidas. No caso de imagens, um procedimento bastante utilizado é o chamado algoritmo de k-médias. O algoritmo denominado Isodata representa um desenvolvimento do algoritmo de k-médias, incorporando procedimentos heurísticos para subdividir ou agrupar aglomerados. Deve-se observar que tanto o algoritmo de k-médias quanto o Isodata têm sido utilizados para quantização tanto no caso unidimensional como no caso vetorial.

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 baseia-se na propriedade de que se duas amostras estão num mesmo agrupamento

num nível do processo elas assim permanecerão até o fim do processo. Os exemplos mais conhecidos de agregação hierárquica ocorrem em Biologia, onde indivíduos se agrupam em espécies, espécies em gêneros, gêneros em família, 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, desenvolve-se atualmente um esforço considerável no sentido de validar estatisticamente estes processos.

É usual em problemas de classificação de padrões adotar a hipótese de que na fase de aprendizado dos parâmetros este estejam 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 esta probabilidade pela classificação do próprio conjunto de amostra de treinamento é considerado otimista, de modo que é necessário dividir o conjunto de amostras numa parte de treinamento e numa parte de teste. Quando se dispõe de um pequeno número de amostras, 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.

AGRADECIMENTOS

O autor deseja agradecer à Comissão Organizadora da Jornada EPUSP/IEEE em Computação Visual pelo honroso convite para ministrar este minicurso e a Janete da Cunha pela cuidadosa datilografia do texto.

BIBLIOGRAFIA

a) Livros

- Chellappa, R.; Sawchuk, A.A. (Eds) Digital Image Processing and Analysis. V.1: Digital Image Processing, New York, NY, IEEE Computer Society, 1985.
- Chellappa, R.; Sawchuk, A.A. (Eds) Digital Image Processing and Analysis. V.2: Digital Image Processing, New York, NY, IEEE Computer Society, 1985.
- Devijver, P.; Kittler, J. Statistical Pattern Recognition, Englewood Cliffs, NJ, Prentice Hall, 1982.
- Devijver, P.; Kittler, J. (Eds) Pattern Recognition Theory and Applications, Berlin, Springer-Verlag, 1987.
- Duda, R.O.; Hart, P.E. Pattern Classification and Scene Analysis, New York, NY, John Willey, 1973.
- Ekstrom, M.P. (Ed.) Digital Image Processing Techniques, New York, NY, Academic, 1984.
- Giardina, C.R.; Dougherty, E.R. Morphological Methods in Image and Signal Processing, Englewood Cliffs, N.J., Prentice Hall, 1988.
- Gonzalez, R.C.; Wintz, P. Digital Image Processing, 2a. Ed., Reading, MA, Addison Wesley, 1987.
- Herman, G.T. Image Reconstruction From Projections - Fundamentals of Computerized Tomography, New York, NY, Academic, 1980.
- Huang, T. (Ed.) Advances in Computer Vision and Image Processing. V.1 - Image Reconstruction from Incomplete Observations, Greenwich, CT, JAI Press, 1984.
- Huang, T.S. (Ed.) Advances in Computer Vision and Image Processing, V.2 - Image Enhancement and Restoration, Greenwich, CT, JAI Press, 1986.
- Huang, T.S. (Ed.) Advances in Computer Vision and Image Processing, V.3 - Time Varying Imagery, Greenwich, CT, JAI Press, 1988.

- Jain, A.K. Fundamentals of Digital Image Processing, Englewood Cliffs, N.J, Prentice Hall, 1989.
- Lim, J.S. Two-Dimensional Signal and Image Processing, Englewood Cliffs, NJ, Prentice Hall, 1990.
- Mascarenhas, N.D.A.; Velasco, F.R.D. Processamento Digital de Imagens, 2a. ed., Kapelus - EBAI, Buenos Aires, 1989.
- Netravali, A.N.; Haskell, B.G. Digital Pictures - Representation and Compression, New York, NY, Plenum, 1988.
- Niemann, H. Pattern Analysis and Understanding, 2a. ed, Berlin, Springer Verlag, 1990.
- Pratt, W.K. Digital Image Processing, New York, NY, John Wiley, 1978.
- Rosenfeld, A. (Ed.) Multiresolution Image Processing and Applications, New York, NY, Springer-Verlag, 1984.
- Rosenfeld, A.; Kak, A.C. Digital Picture Processing, 2a. ed, New York, NY, Academic, 1982.
- Schalkoff, R.J. Digital Image Processing and Computer Vision, New York, NY, John Wiley, 1989.
- Stark, H. (Ed.) Image Recovery: Theory and Applications, New York, NY, Academic, 1987.
- Therrien, C.W. Decision Estimation and Classification - An Introduction to Pattern Recognition and Related Topics, New York, NY, John Wiley, 1989.
- Young, T.Y.; Fu, K.S. (Eds.) Handbook of Pattern Recognition and Image Processing, New York, NY, Academic, 1986.

b) Revistas

APPLIED OPTICS, Washington, D.C., Optical Society of América, 1962 -, quinz.

COMPUTER VISION, GRAPHICS AND IMAGE PROCESSING. New York N.Y., Academic, 1972 -, mens.

IEEE TRANSACTIONS ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, New York, N.Y., IEEE Signal Processing Society, 1974 -, mens.

IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, New York, N.Y., IEEE Circuits and Systems Society, 1974 -, mens.

IEEE TRANSACTIONS ON MEDICAL IMAGING, New York, N.Y., IEEE Signal Processing Society, 1982 -, mens.

IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, New York, N.Y., IEEE Systems, Man and Cybernetics Society, 1971, -, bim.

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

IMAGE AND VISION COMPUTING, London, Butterworths, 1982 -, trim.

INTERNATIONAL JOURNAL OF PATTERN RECOGNITION AND ARTIFICIAL INTELLIGENCE, Singapore, World Scientific, 1987 -, trim.

JOURNAL OF THE OPTICAL SOCIETY OF AMERICA, PART A-OPTICS AND IMAGE SCIENCE, Washington, D.C., Optical Society of America, 1984 -, mens.

OPTICAL ENGINEERING, Bellingham, WA, The International Society for Optical Engineering, 1962 -, mens.

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

PATTERN RECOGNITION LETTERS, Amsterdam, North Holland, 1982 -, mens.

PROCEEDINGS OF THE IEEE, New York, N.Y., The Institute of Electrical and Electronics Engineers, 1963 -, mens.

SIGNAL PROCESSING, Amsterdam, Elsevier, 1979 -, mens.

c) Anais de Conferências

IEEE Conference on Computer Vision and Pattern Recognition 83, 85, 86, 88, 89
- IEEE Computer Society Press.

IEEE Conference on Pattern Recognition and Image Processing 77, 78, 79, 81, 82
- IEEE Computer Society Press.

International Conference on Acoustics, Speech and Signal Processing - anual
desde 76 - IEEE Signal Processing Society.

International Conference on Computer Vision- 87, 88, 90- IEEE Computer Society
Press.

International Conference on Pattern Recognition, (até 1978, International
Joint Conference on Pattern Recognition) 73, 74 76, 78, 80, 82, 84, 86, 88,
90 - IEEE Computer Society Press.

d) Números Especiais Recentes sobre Processamento de Imagens

IEEE Journal on Selected Areas in Communications-Special Issue on Low Bit-Rate
Coding of Moving Images, 5(7), Agosto 1987.

IEEE Transactions on Acoustics, Speech and Signal Processing, Special Section
on Multidimensional Signal Processing in Image Analysis, 37(12), Dezembro
1989.

IEEE Transactions on Circuits and Systems, Special Issue on Digital Image
Processing and Applications 34(11), Novembro 1987.

IEEE Transactions on Circuits and Systems, Special Issue on VLSI
Implementation for Digital Image and Video Processing Applications, 36(11),
Outubro 1989.

IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issues on Industrial Machine Vision and Computer Vision Technology, Part I, 10(1), Janeiro 1988, Part II, 10(3), Maio 1988.

IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Section on Computer Architectures, 11(3), Março 1989.

IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Issue on Visual Motion, 11(5), Maio 1989.

IEEE Transactions on Pattern Analysis and Machine Intelligence, Special Section on Multiresolution Representation, 11(7), Julho 1989.

IEEE Transactions on Systems, Man and Cybernetics, Special Issue on Computer Vision, 19(6), Novembro/Dezembro 1989.

Journal of Applied Statistics - Special Issues on Image Processing, 16(2 e 3), 1989.

Optical Engineering - Special Issue on Visual Communications and Image Processing, 26(7), Julho 1987.

Optical Engineering - Special Issue on Visual Communications and Image Processing, 28(7), Julho 1989.

Optical Engineering - Special Issue on Image Restoration and Reconstruction, 29(7), Maio 1990.

Proceedings of the IEEE - Special Issue on Visual Communications Systems, 73(4), Abril 1985.

Proceedings of the IEEE - Special Issue on Perceiving Earth's Resources from Space, 73(6), Junho 1985.

Proceedings of the IEEE - Special Issue on Computer Vision, 76(8), Agosto 1988.

Proceedings of the IEEE - Special Issue on Multidimensional Signal Processing, 78(4), Abril 1990.

Signal Processing - Special Issue on Multidimensional Signal Processing - Part 1, 15(3), Outubro 1988.

Signal Processing - Special Issue on Multidimensional Signal Processing - Part 2, 16(1), Janeiro 1989.

Signal Processing - Special Issue on Advances in Mathematical Morphology, 16(4), Abril 1989.