

Imprimir

Fechar

Referência Completa

Tipo da Referência Conference Proceedings

Repositório sid.inpe.br/iris@1912/2005/07.18.23.35.47

Metadados sid.inpe.br/iris@1912/2005/07.18.23.35.56

Site mtc-m12.sid.inpe.br

Rótulo 528

Chave Secundária INPE-3325-PRE/624

Chave de Citação MascarenhasVela:1984:PrDiIm

Autor 1 Mascarenhas, Nelson Delfino D'Avila
2 Velasco, Flavio Roberto Dias

Grupo 1 DPI-INPE-BR

Título Processamento digital de imagens 

Nome do Evento Escola de Computação, 4.

Ano 1984

Data 12-20 jul. 1984

Localização do Evento Sao Paulo, BR

Palavras-Chave RECONHECIMENTO DE PADROES, PROCESSAMENTO DIGITAL DE SINAIS, ANALISE DE IMAGENS, DIGITAL PROCESSING AND CORRECTION, IMAGE PROCESSING, PATTERN RECOGNITION, DIGITAL SIGNAL PROCESSING, IMAGE ANALYSIS.

Resumo Este texto foi preparado para servir de base para um minicurso em processamento digital de imagens oferecido na Escola de Computacao, Instituto de Matematica e Estatistica, da Universidade de Sao Paulo - IME/USP, cobrindo os principais topicos da area.

Páginas 315

Idioma Pt

Tipo Secundário PRE CN

Divulgação <E>

Area SRE

Projeto PROCIM

Ultima Atualização dos Metadados 2015:03.18.16.27.57 sid.inpe.br/bibdigital@80/2006/04.07.15.50 administrator

Estágio do Documento concluido

e-Mail (login) marciana

Grupo de Usuários administrator

Visibilidade shown

Transferível 1

Tipo do Conteúdo External Contribution

Unidade Imediatamente Superior 8JMKD3MGPCW/3EQCCU5

Conteúdo da Pasta source não têm arquivos

Conteúdo da Pasta agreement não têm arquivos

Histórico 2008-06-09 19:43:37 :: administrator -> jefferson ::
2010-07-07 18:43:40 :: jefferson -> administrator ::
2015-03-18 16:27:57 :: administrator -> marciana :: 1984

Campos Vazios accessionnumber affiliation archivingpolicy archivist booktitle callnumber copyholder copyright creatorhistory descriptionlevel documentstage doi e-mailaddress edition editor electronicmailaddress format isbn issn lineage mark mirrorrepository nextedition notes numberoffiles numberofvolumes organization parameterlist parentrepositories previousedition progress publisher publisheraddress readergroup readergroup readpermission resumeid rightsholder secondarydate secondarymark serieseditor session shorttitle size sponsor subject targetfile tertiarymark tertiarytype type url versiontype volume

1. Publicação nº <i>INPE-3325-PRE/624</i>	2. Versão	3. Data <i>Outubro 1984</i>	5. Distribuição <input type="checkbox"/> Interna <input checked="" type="checkbox"/> Externa <input type="checkbox"/> Restrita
4. Origem <i>DPI</i>	Programa <i>PROCIM</i>		
6. Palavras chaves - selecionadas pelo(s) autor(es) <i>Processamento de Imagens Processamento Digital de Sinais</i> <i>Classificação de Padrões Análise de Imagens</i>			
7. C.D.U.: <i>621.376.5</i>			
8. Título <i>PROCESSAMENTO DIGITAL DE IMAGENS</i>		<i>INPE-3325-PRE/624</i>	10. Páginas: <i>315</i>
			11. Última página: <i>A.18</i>
9. Autoria <i>Nelson D.A.Mascarenhas</i> <i>Flávio Roberto D.Velasco</i>			12. Revisada por <i>G. Gyan</i> Membros do DIN/DPI
Assinatura responsável <i>Nelson Mascarenhas</i>			13. Autorizada por <i>Parada</i> Nelson de Jesus Parada Diretor Geral
14. Resumo/Notas <i>Este texto foi preparado para servir de base para um minicurso em processamento digital de imagens oferecido na Escola de Computação, Instituto de Matemática e Estatística, da Universidade de São Paulo - IME/USP, cobrindo os principais tópicos da área.</i>			
15. Observações <i>Texto apresentado na 4ª Escola de Computação, IME/USP, que se realizou de 12 a 20 de julho de 1984, em São Paulo, SP.</i>			

ABSTRACT

This text was prepared as a basis for a short course on digital image processing, offered at the school of Computation, IME-USP, São Paulo, July 1984, covering the main topics of the area.

"PROCESSAMENTO DIGITAL DE IMAGENS"

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

Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq
Instituto de Pesquisas Espaciais - INPE

Flávio Roberto Dias Velasco é pesquisador no Departamento de Informática do Instituto de Pesquisas Espaciais (INPE/CNPq), São José dos Campos, SP. Formou-se em Engenharia Eletrônica no Instituto Tecnológico de Aeronáutica em 1970, obteve o grau de Mestre em Sistemas e Computação pela COPPE/UFRJ e o grau de Doutor em Computação Aplicada pelo INPE/CNPq, em 1977. Suas áreas de interesse incluem processamento de imagens, visão por computador, engenharia de software e software básico.

Nelson Delfino d'Ávila Mascarenhas é Pesquisador no Departamento de Informática do Instituto de Pesquisas Espaciais (INPE/CNPq), São José dos Campos, SP. Formou-se em Engenharia Eletrônica pelo ITA em 1966, obteve o grau de Mestre em Ciências em Engenharia Eletrônica pelo mesmo Instituto em 1969 e o doutoramento em Engenharia Elétrica pela University of Southern California em 1974. Suas áreas de interesse incluem processamento de imagens, reconhecimento de padrões, processamento digital de sinais e teoria estatística de comunicações.

PREFÁCIO

A área de processamento digital de imagens teve início em meados da década de 60, com a motivação criada pelos programas espaciais da NASA, nos EUA. O aumento da capacidade de memória e velocidade dos computadores contribuiu para dar impulso à nova tecnologia, que passou a ser aplicada em outras áreas tais como: medicina, microscopia, ótica ou eletrônica, transmissão digital de sinais de televisão ou fac-símile, e, recentemente, em sensores visuais de robôs industriais.

Na área espacial deve ser mencionada a grande utilização de imagens digitais a partir do lançamento da série de satélites LANDSAT em 1972, bem como dos satélites meteorológicos, tanto de órbita polar como geostacionária. Em 1973, a invenção do tomógrafo computadorizado por G. N. Hounsfield abriu amplas perspectivas de utilização das técnicas digitais em imagens radiográficas.

Para o futuro surgem novas possibilidades: o lançamento de novas séries de satélites de observação da terra (inclusive os nacionais, previstos para o início da próxima década); novas aplicações médicas (como tomografia por ressonância magnética nuclear ou por emissão de pósitrons e radiografia digital); aumento das aplicações industriais de robótica; uso de sistemas de teleconferência e fac-símile, etc.

Processamento Digital de Imagens é essencialmente uma área multidisciplinar, derivando suas técnicas de Ótica, Colorimetria, Neurofisiologia, Engenharia Elétrica e Ciência de Computação. O texto aqui apresentado concentrou sua atenção nessas duas últimas disciplinas. Não foram abordados também aspectos de "hardware" como sensores, visualizadores ou arquiteturas especiais.

A despeito de se ter feito um esforço para tornar o texto completo, a escolha do material em cada um dos capítulos reflete mais a perspectiva pessoal dos autores sobre os respectivos tópicos, do que a importância em relação a outros tópicos não abordados.

São José dos Campos, maio de 1984.

AGRADECIMENTOS

O autores desejam agradecer: à Comissão Organizadora da 4ª Escola da Computação a oportunidade de ministrar este curso; ao Instituto de Pesquisas Espaciais (INPE/CNPq) o constante apoio oferecido durante a elaboração do texto; ao pessoal do Laboratório de Tratamento de Imagens Digitais do INPE a assistência na digitalização e processamento de imagens; a José Antonio G. Pereira, Ubirajara M.B. Lima, Gilberto Câmara Neto, Guaraci José Erthal, Virgínia R. Moraes Correia, Edson L.F. Senne, Neusa Maria D. Bicudo, Glória C. Bertti, Maria do Carmo S. Soares e Marília P. de Carvalho a revisão do texto; a Antonio Carnevale Neto, Guaraci José Erthal e Gilberto Câmara Neto, o apoio computacional; a Beatriz M.T.Z. Parreiras o trabalho fotográfico de ilustrações do texto; a Idelfonso de Oliveira Filho e Renêa Pavanelli os desenhos utilizados; a Tânia R.F. Sanchez e Mônica A. de Oliveira o trabalho datilográfico; a Marciana Leite Ribeiro a revisão das referências bibliográficas.

Além disso, os autores desejam estender seus agradecimentos às seguintes pessoas e instituições que fornecerem imagens utilizadas no texto: R.C. Gonzalez e P. Wintz (Figuras 2.1 e 2.2 do livro "Digital Image Processing", Addison Wesley, 1977); T.S. Huang (Figura 2.9 de "PCM Picture Transmission", IEEE Spectrum, vol. 2, nº 12, 1965, pp. 57-63); T. S. Huang, O.J. Tretiak, B. Prasada e Y. Yamaguchi (Figura 2.12 de "Design Considerations in PCM Transmission of Low Resolution Monochrome Still Pictures", Proc. IEEE, vol. 55, 1967, pp. 331-335); B. R. Hunt (Figura 4.4 de "Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer", IEEE Trans. Computers, vol. c-22, nº 9, 1973, pp. 805-812); Ubirajara M.B. Lima (Figuras 6.12, 6.13 e 6.14 de "Compressão de Imagens Digitais por Transformação", Relatório INPE-2893-PRE/421, setembro 1983); Eni A. Oliveira (Figuras 3.3, 8.5a, 8.5b); Optical Society of America e Dr. B.L. Mc. Glamery (Figura 4.3 de "Restoration of Turbulence-Degraded Images") J. Opt. Soc. Amer., vol. 57, nº 3, 1967, pp. 293-297); José A. G. Pereira (Figuras 7.2, 7.3 e 7.4 de "Métodos de Registro de Imagens Usando Técnicas de Testes Sequenciais de Hipóteses", Relatório INPE-2681-TDL/124, Março 1983); José A.G. Pereira; Guaraci J. Erthal (Figura 2.10); William K. Pratt (Figuras 2.5; 2.11; 6.8; 6.9; 6.10; 6.11; 6.15 e 6.16 do livro "Digital Image Processing", John Wiley, 1978); Projeto Portinari (Figura 2.6); Projeto Portinari (Figura 2.6); Unidade Radiológica Paulista (Figura 5.1).

SUMÁRIO

	<u>Pág.</u>
<u>CAPÍTULO 1 - INTRODUÇÃO</u>	
1.1 - Imagens	1.1
1.2 - Sistemas de imageamento	1.2
1.3 - Processamento digital de imagens	1.4
REFERÊNCIAS BIBLIOGRÁFICAS	1.8
FIGURAS	1.10
<u>CAPÍTULO 2 - DIGITALIZAÇÃO</u>	
2.1 - Amostragem	2.1
2.1.1 - Caso unidimensional	2.1
2.1.2 - Caso bidimensional	2.2
2.2 - Quantização	2.4
2.2.1 - Quantização ótima	2.4
2.2.2 - Aproximação da quantização ótima	2.7
2.2.3 - Falsos contornos	2.8
2.2.4 - Quantização visual	2.8
2.2.5 - Quantização por ruído pseudoaleatório	2.9
2.2.6 - Relação entre amostragem e quantização	2.10
REFERÊNCIAS BIBLIOGRÁFICAS	2.11
FIGURAS	2.12
<u>CAPÍTULO 3 - REALCE</u>	
3.1 - Transformação da escala de cinza	3.2
3.2 - Modificação de histograma	3.3
3.3 - Filtragem	3.6
3.3.1 - Filtragem por convolução	3.9
3.3.2 - Filtragem não-linear	3.17
REFERÊNCIAS BIBLIOGRÁFICAS	3.21
FIGURAS	3.23
<u>CAPÍTULO 4 - RESTAURAÇÃO</u>	
4.1 - Modelos de degradação	4.1
4.2 - Determinação experimental da função de espalhamento pontual ...	4.4
4.2.1 - Fonte pontual	4.4
4.2.2 - Fonte retilínea	4.5
4.2.3 - Fonte de borda retilínea	4.7

	<u>Pág.</u>
4.3 - Filtragem inversa	4.8
4.4 - O problema do malcondicionamento e sua solução	4.9
4.4.1 - Estimador de mínimos quadrados	4.10
4.4.2 - Estimador de mínimos quadrados com restrição linear de desigualdade - conhecimento a priori determinístico	4.13
4.4.3 - Estimador de mínimos quadrados com restrição quadrática de igualdade - técnicas de suavização e regularização	4.13
4.4.4 - Filtragem de Wiener - conhecimento a priori estatístico	4.16
4.4.4.1 - Caso contínuo	4.17
4.4.4.2 - Caso discreto	4.19
4.5 - Problema bidimensional - solução no domínio de Fourier	4.21
4.6 - Outras técnicas de restauração	4.24
REFERÊNCIAS BIBLIOGRÁFICAS	4.26
FIGURAS	4.29
 CAPÍTULO 5 - RECONSTRUÇÃO	
5.1 - Teorema do corte de Fourier	5.3
5.2 - Algoritmo de filtragem e retroprojeção	5.6
5.3 - Técnicas algébricas	5.9
5.3.1 - O método de Kaczmarz	5.9
5.3.2 - Utilização de desigualdades lineares	5.12
5.3.3 - Algoritmo de mínimos quadrados - métodos de suavização - métodos bayesianos	5.14
5.3.4 - Otimização da entropia	5.15
5.4 - Comparação entre métodos - problemas abertos	5.16
REFERÊNCIAS BIBLIOGRÁFICAS	5.18
FIGURAS	5.21
 CAPÍTULO 6 - CODIFICAÇÃO	
6.1 - O problema de codificação da fonte de sinal - Teoria de razão de distorção	6.2
6.2 - Técnicas reversíveis	6.5
6.2.1 - Codificação de um elemento de imagem	6.6
6.2.2 - Codificação de elementos adjacentes	6.8
6.3 - Técnicas irreversíveis	6.8
6.3.1 - Modelos matemáticos	6.8
6.3.2 - Técnicas preditivas	6.9

	<u>Pág.</u>
6.3.3 - Técnicas por transformação	6.14
6.3.3.1 - Transformação ótima	6.14
6.3.3.2 - Transformações sub-ótimas	6.15
6.3.3.3 - Seleção dos coeficientes	6.17
6.3.3.4 - Quantização e codificação dos coeficientes	6.18
6.3.3.5 - Efeitos de erros de transmissão	6.18
6.3.4 - Codificação híbrida	6.19
6.3.5 - Outros métodos de codificação	6.19
REFERÊNCIAS BIBLIOGRÁFICAS	6.21
FIGURAS	6.23
 <u>CAPÍTULO 7 - REGISTRO</u>	
7.1 - Registro translacional	7.2
7.1.1 - Registro translacional de imagens binárias	7.6
7.2 - Transformações geométricas	7.7
7.2.1 - Determinação da função de transformação	7.8
7.2.2 - Métodos de interpolação de imagens	7.13
REFERÊNCIAS BIBLIOGRÁFICAS	7.19
FIGURAS	7.21
 <u>CAPÍTULO 8 - SEGMENTAÇÃO</u>	
8.1 - Detecção de borda	8.1
8.1.1 - Perseguição de contorno	8.3
8.1.2 - Pesquisa em grafo	8.5
8.1.3 - Transformada de Hough	8.8
8.2 - Detecção de regiões	8.9
8.2.1 - Crescimento de regiões	8.11
8.2.2 - Método da divisão e fusão	8.13
8.2.3 - Segmentação por classificação de pontos	8.15
REFERÊNCIAS BIBLIOGRÁFICAS	8.18
FIGURAS	8.21
 <u>CAPÍTULO 9 - REPRESENTAÇÃO E DESCRIÇÃO</u>	
9.1 - Esquemas de representação	9.2
9.1.1 - Representação por varredura	9.2
9.1.2 - Representação de borda	9.3
9.1.3 - Representação por árvores quaternárias ("quadrees")	9.4

	<u>Pág.</u>
9.1.4 - Conversão entre representações	9.5
9.2 - Operações com imagens	9.9
9.2.1 - Encolhimento e expansão	9.10
9.2.2 - Operações lógicas em regiões	9.10
9.3 - Propriedades geométricas	9.12
9.3.1 - Tamanho	9.13
9.3.1.1 - Área	9.13
9.3.1.2 - Perímetro e comprimento de arco	9.15
9.3.1.3 - Extensão, seção cruzada e assinatura	9.16
9.3.2 - Forma	9.16
9.3.2.1 - Complexidade	9.16
9.3.2.2 - Alongamento	9.17
9.3.2.3 - Convexidade	9.17
REFERÊNCIAS BIBLIOGRÁFICAS	9.19
FIGURAS	9.21
 <u>CAPÍTULO 10 - CLASSIFICAÇÃO DE IMAGENS</u>	
10.1 - Noções básicas de classificação	10.1
10.2 - Extração de atributos	10.3
10.2.1 - Transformações no espaço de atributos	10.3
10.2.2 - Medidas de informação e distância	10.7
10.2.2.1 - Divergência	10.8
10.2.2.2 - Distância J-M	10.11
10.3 - Elementos de teoria estatística de decisão	10.13
10.4 - Treinamento e aprendizado	10.18
10.4.1 - Aprendizado com supervisão	10.18
10.4.2 - Aprendizado sem supervisão	10.20
10.4.3 - Técnicas não paramétricas	10.21
10.4.4 - Algoritmos de treinamento	10.21
10.4.5 - Técnicas de agregação de dados	10.25
10.5 - Avaliação de desempenho do classificador	10.28
10.6 - Desenvolvimentos recentes e tendências futuras	10.30
REFERÊNCIAS BIBLIOGRÁFICAS	10.32
FIGURAS	10.36

APÊNDICE

A.1 - Sistemas lineares	A.1
A.2 - Análise de Fourier	A.3
A.3 - Transformações discretas em imagens	A.5
A.3.1 - Transformada discreta de Fourier	A.7
A.3.2 - Transformada discreta do cosseno	A.10
A.3.3 - Transformada discreta de Hadamard	A.11
A.3.4 - Transformação discreta de Karhunen-Loeve	A.11
A.4 - Noções de campos aleatórios	A.15
REFERÊNCIAS BIBLIOGRÁFICAS	A.18

CAPÍTULO 2

DIGITALIZAÇÃO

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, é necessário discretizar a imagem tanto no espaço quanto na amplitude. A discretização no espaço é denominada amostragem, enquanto a discretização na amplitude é chamada quantização.

2.1 - AMOSTRAGEM

2.1.1 - CASO UNIDIMENSIONAL

A ferramenta fundamental para determinar a mínima razão de amostragem é dada pelo teorema de Whittaker-Kotelnikov-Shannon (Rosenfeld and Kak, 1982). Em sua forma unidimensional, o teorema pode ser expresso da seguinte maneira.

TEOREMA 2.1 - *Se a Transformada de Fourier de uma função $f(t)$ anula-se para $|u| \geq W$, então $f(t)$ pode ser reconstruída exatamente por amostras com uma separação de $\frac{1}{2W}$.*

Essa separação mínima é também chamada razão de Nyquist. Uma descrição sumária desse teorema pode ser feita na seguinte forma (Veja-se a Figura 2.1): o processo de amostragem ideal de uma função pode ser representado pela multiplicação dessa função por um trem de impulsos (funções de Dirac) separados pela taxa de amostragem. A Transformada de Fourier do produto vai ser dada pela convolução das Transformadas da função original e do trem de impulsos. Disso resulta que o espectro da função amostrada é a repetição periódica do espectro da função original. Se a amostragem é suficientemente fina, essa repetição não vai implicar uma superposição dos espectros no domínio da frequência. A imagem original pode então ser recuperada da imagem amostrada por um filtro ideal de espectro retangular que deixa passar apenas as frequências correspondentes às da imagem original. A multiplicação do espectro do filtro ideal com o espectro repetitivo da função amostrada corres

ponde à convolução da função amostrada com uma função do tipo $\frac{\text{sen } x}{x}$. Pode ser mostrado (Sakrison, 1968) que isso equivale a um processo de interpolação das amostras originais por essa última função, segundo a expressão:

$$f(x) = \sum_{K=-\infty}^{+\infty} f\left(\frac{K}{2W}\right) \frac{\text{sen} \left[2\pi W \left(x - \frac{K}{2W} \right) \right]}{2\pi W \left(x - \frac{K}{2W} \right)}. \quad (2.1)$$

A Figura 2.1 mostra também a situação em que a amostragem não é suficientemente fina. Nesse caso vai haver uma superposição dos espectros replicados e não vai ser possível reconstruir o sinal original pelo filtro passa-baixas ideal, o que é denominado fenômeno de "aliasing".

A situação tratada até aqui supõe a existência de um trem infinito de impulsos para efetuar a amostragem. Numa abordagem mais realista o fato de o número de impulsos ser finito é levado em conta (Figura 2.2) pela multiplicação por uma função do tipo pulso da sequência infinita de impulsos original. Isso vai corresponder no domínio da frequência a uma convolução do espectro periódico da função amostrada idealmente com uma função do tipo sinc u, o que vai provocar uma distorção adicional do sinal a ser recuperado pelo filtro passa-baixas ideal. Assim sendo, em geral não é possível recuperar a função original mesmo que as condições do teorema da amostragem sejam satisfeitas.

2.1.2 - Caso Bidimensional

Os conceitos apresentados para sinais unidimensionais podem ser adaptados aos sinais bidimensionais, que podem ser utilizados para representar imagens. Nesse caso, a amostragem ideal é representada por uma matriz infinita de funções delta de Dirac numa grade de espaçamento Δx na direção horizontal e Δy na direção vertical, o que pode ser expresso por:

$$s(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x-i\Delta x, y-j\Delta y) \quad (2.2)$$

e está representada graficamente na Figura 2.3.

A imagem contínua original, cujo espectro está mostrado na Figura 2.4a, pelo processo de amostragem, é multiplicada pela matriz de impulsos e, pelo processo de convolução no domínio da frequência, isso resulta em uma repetição periódica em duas dimensões do espectro da função original (Figura 2.4b), com separação $\frac{1}{\Delta x}$ no eixo u (horizontal) e $\frac{1}{\Delta y}$ no eixo v (vertical). Se o espectro da imagem original for limitado em frequência e o espaçamento do processo de amostragem for suficientemente fino não haverá superposição no espectro repetitivo da função amostrada e a imagem original poderá ser recuperada por uma filtragem passa-baixas ideal que selecionará a porção do espectro centralizada na origem do plano u-v de frequências.

Se a amostragem não é suficientemente fina para que as condições anteriores sejam satisfeitas ocorrerá o efeito de "aliasing" que, no caso de imagens, também é conhecido como fenômeno de Moiré ou padrões de Moiré. A Figura 2.5 ilustra esse efeito para uma imagem de teste.

Na prática, imagens reais tendem a conter componentes de alta frequência devido a bordas ou mesmo à presença de ruído. Assim sendo, sempre ocorrerá em alguma escala o fenômeno de "aliasing". Todavia o efeito visual desse fenômeno é negligenciável se 90% a 95% da energia do espectro da imagem original é contida em frequências que satisfazem o teorema da amostragem (Moik, 1980). Além disso, os efeitos de "aliasing" são reduzidos por sistemas reais de amostragem, onde a função amostradora não é exatamente uma matriz de funções delta de Dirac.

A Figura 2.6 ilustra o efeito visual de subamostragem por replicação de elementos obtidos pela amostragem original. O efeito de blocos é diminuído se a imagem for observada a uma certa distância, o que equivale a efetuar uma filtragem passa-baixas devido à limitação em frequência espacial do sistema visual humano.

É possível formular uma versão do teorema da amostragem na situação em que os sinais são descritos por processos aleatórios (caso unidi-

mensional) ou campos aleatórios (caso bidimensional). Assim sendo, se a densidade espectral de potência anula-se fora de uma região limitada, é possível reproduzir o sinal aleatório por amostras com erro médio quadrático nulo se as amostras forem suficientemente próximas. Neste caso, a função interpoladora reproduzirá exatamente a função de correlação do sinal aleatório a partir das amostras dessa função de correlação.

2.2 - QUANTIZAÇÃO

Pelo processo de quantização, uma imagem com tons contínuos é convertida numa com tons discretos. Para armazenamento ou transmissão digital, cada tom é representado por um código binário. A escolha desse código de maneira a minimizar a taxa de bits necessária ao armazenamento ou à transmissão da imagem será discutida na seção sobre codificação.

2.2.1 - Quantização ótima

Os métodos de quantização buscam, 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 originais é suposta conhecida, por exemplo $f_x(x)$. Também é suposto que o escalar x que representa o brilho da imagem seja restrito ao intervalo $X_I \leq x \leq X_S$.

O problema de quantização consiste na especificação de níveis de decisão $D_0 \equiv X_I, D_1, \dots, D_N \equiv X_S$ e um conjunto de níveis de reconstrução R_0, R_1, \dots, R_{N-1} tal que se $D_i \leq x \leq D_{i+1}$ então a amostra é atribuída ao nível de reconstrução R_i (Figura 2.7).

O erro médio quadrático pode ser expresso por:

$$\begin{aligned} \epsilon &= E \left\{ (x - \bar{x})^2 \right\} = \int_{x_I}^{x_S} (x - \bar{x})^2 f_x(x) dx = \\ &= \sum_{j=0}^{N-1} \int_{D_j}^{D_{j+1}} (x - R_j)^2 f_x(x) dx \end{aligned} \quad (2.3)$$

Igualando a zero as derivadas parciais de ϵ com respeito aos níveis de decisão D_i e reconstrução R_i têm-se:

$$\frac{\partial \epsilon}{\partial D_i} = 0 = (D_{i+1} - R_i)^2 f_x(D_{i+1}) - (D_i - R_{i-1})^2 f_x(D_i). \quad (2.4)$$

$$\frac{\partial \epsilon}{\partial R_i} = 0 = 2 \int_{D_i}^{D_{i+1}} (x - R_i) f_x(x) dx. \quad (2.5)$$

Assim sendo, obtêm-se o seguinte conjunto de equações não-lineares simultâneas:

$$D_i = \frac{R_{i-1} + R_i}{2}, \quad (2.6)$$

$$R_i = \frac{\int_{D_i}^{D_{i+1}} x f_x(x) dx}{\int_{D_i}^{D_{i+1}} f_x(x) dx} = E[x/x \in (D_i, D_{i+1})]. \quad (2.7)$$

O resultado anterior indica que os níveis de decisão devem ser o ponto médio dos níveis de reconstrução e que estes, por sua vez, devem ser o centro de massa da distribuição dentro dos intervalos de decisão.

Quando a densidade \bar{e} é uniforme, a quantização também é uniforme, ou seja:

$$D_i = \frac{R_{i+1} + R_i}{2}, \quad (2.8)$$

$$\hat{R}_i = \frac{D_i + D_{i+1}}{2}. \quad (2.9)$$

Quando a densidade de probabilidade \bar{e} é gaussiana, Max (1960) obteve a solução para o conjunto de equações não-lineares e computou tabelas de níveis ótimos de decisão e reconstrução como função do número de níveis de quantização.

Quando as Equações 2.6 e 2.7 são satisfeitas pode ser mostrado (Pratt, 1978) que o erro médio quadrático mínimo pode ser dado por:

$$\epsilon_{\min} = E\{x^2\} - \sum_{j=0}^{N-1} R_j^2 \cdot P \{D_j \leq x \leq D_{j+1}\} \quad (2.10)$$

Além disso, os seguintes resultados de interesse teórico são válidos (Jain, 1981):

1) A saída do quantizador \bar{x} é um estimador sem viés da entrada, isto é:

$$E\{\bar{x}\} = E\{x\}. \quad (2.11)$$

2) O erro de quantização \bar{x} é ortogonal à saída de quantizador, isto é:

$$E\{x - \bar{x}\} \cdot \bar{x} = 0. \quad (2.12)$$

Essa última relação fornece um modelo para o quantizador na forma:

$$x = \bar{x} + \eta, \quad (2.13)$$

$$E\{x_n\} = E\{\eta^2\}, \quad (2.14)$$

onde η representa o erro de quantização que é não-correlacionado com a saída do quantizador.

2.2.2 - Aproximação da quantização ótima

Quando o número de níveis de quantização N é grande, a densidade de probabilidade $f_x(x)$ pode ser considerada constante sobre cada intervalo de quantização. Nessas condições, minimizando ϵ com respeito aos níveis de reconstrução R_j , tem-se:

$$R_j = \frac{D_{j+1} + D_j}{2}. \quad (2.15)$$

Por sua vez, a localização ótima dos níveis de decisão foi obtida por Panter e Dite (1951) pelo método dos multiplicadores de Lagrange, daí resultando aproximadamente em:

$$D_j = X_I + (X_S - X_I) \cdot \frac{\int_{X_I}^{X_j} [f_x(x)]^{-1/3} dx}{\int_{X_I}^{X_S} [f_x(x)]^{-1/3} dx} \quad (2.16)$$

$$\text{onde } X_j = X_I + \frac{j}{N} (X_S - X_I). \quad (2.17)$$

Se a distribuição é uniforme, também os níveis de decisão D_j serão o ponto médio dos níveis de reconstrução, o que implica uma quantização uniforme. Se a distribuição não é uniforme, tanto o resultado ótimo de Max (1960) quanto a aproximação de Panter e Dite (1951) implicam um menor espaçamento dos níveis de decisão para grandes valores da densidade de probabilidade. Essa quantização não-uniforme dos tons de cinza pode ser efetuada por uma sucessão de três operações, como mostra a Figura 2:8.

A primeira operação é tal que, na sua saída, a variável aleatória u é uniformemente distribuída entre $-1/2$ e $+1/2$.

Pode-se demonstrar (Davenport, 1970) que a função que realiza essa operação é dada por:

$$u = F_x(x) - 1/2, \quad (2.18)$$

onde $F_x(x)$ é a função de distribuição da variável aleatória de entrada do processo. Essa função efetua uma compressão dos dados de entrada, que é seguida de uma quantização uniforme e de uma expansão na saída obtida pela função inversa daquela dada por 2.11.

A Figura 2.9 ilustra o efeito da quantização não-uniforme com 16 e 4 níveis de cinza.

2.2.3 - Falsos contornos

A redução do número de tons de cinza em imagens monocromáticas, pela diminuição do número de níveis de quantização, conduz ao fenômeno denominado falsos contornos. Isso se caracteriza pelo aparecimento de transições visíveis de tons de cinza em regiões com variação relativamente suave de tonalidade. Para assegurar que falsos contornos não apareçam numa imagem monocromática são necessários normalmente 64 níveis de cinza (quantização uniforme com 6 bits). A Figura 2.10 ilustra o efeito visual da redução do número de níveis de quantização.

Quantizadores ótimos segundo o critério de mínimo erro médio casados ao histograma da imagem podem eliminar os falsos contornos com apenas 32 níveis (5 bits).

2.2.4 - Quantização visual

É possível realizar um esquema de quantização do tipo expansão-quantização uniforme - compressão, do tipo ilustrado pela Figura 2.8, levando em conta propriedades de percepção visual humana. Se L e $L+\Delta L$ são dois va

lores de luminância tal que ΔL é o menor valor incremental perceptível, há resultados experimentais que indicam que $\frac{\Delta L}{L}$ é aproximadamente constante para uma ampla gama de valores de L . Esta é a chamada Lei de Weber - Fechner (Hecht, 1924). Se se definir o contraste C pela relação:

$$C = \log L, \quad (2.19)$$

então $\Delta C = \frac{\Delta L}{L} \approx$ constante. Isto significa que a sensibilidade visual é aproximadamente uniforme a variações de C . Assim sendo, têm sido propostas na literatura (Jain, 1981; Kretz, 1975) funções do tipo logaritmo (por exemplo $\alpha \log(1+\beta L)$, $L \in (0,1)$, onde α e β são constantes) para o primeiro bloco da Figura 2.8. Resultados experimentais indicam que com a quantização do contraste, 4 bits são suficientes para praticamente eliminar os falsos contornos.

2.2.5 - Quantização por Ruído Pseudoaleatório

Este método, desenvolvido por Roberts (1962), permite minimizar o efeito de falsos contornos pela adição de um ruído pseudoaleatório com pequena variância antes da quantização e subtraindo-o sincronamente após a quantização, para a exibição da imagem. Adicionando-se o ruído com média nula, algumas amostras a serem quantizadas superam os níveis de decisão originais, enquanto outros passam a ser inferiores a esses níveis. Em geral, o nível de ruído deve ser escolhido de modo a afetar o bit menos significativo do quantizador. Na exibição da imagem, o ruído tendo a preencher as regiões de contornos, de modo que a média da imagem não se altera. Todavia, há um aumento do erro médio quadrático e uma degradação do tipo "neve" é introduzida na imagem. De qualquer modo, é possível realizar uma quantização até com 3 bits sem um efeito visível de contornos. A Figura 2.11 ilustra o efeito da adição de ruído pseudoaleatório na quantização de uma imagem.

2.2.6 - Relação entre Amostragem e Quantização

Dado o número total de bits a ser utilizado para representar uma imagem o problema de escolher o número de amostras e o número de níveis de quantização é bastante dependente do tipo de imagem e de critérios de julgamento subjetivo. A Figura 2.12 ilustra essa dependência.

Considerando a Figura 2.12 (a e b), imagens de uma garota com distribuição relativamente suave de tons de cinza, um observador normalmente iria preferir a representação com menor número de elementos e maior número de níveis de quantização (Figura 2.12b). Já no caso da imagem da multidão (Figura 2.12c e d), de variação rápida de tons de cinza, o oposto tende a prevalecer. Huang et al (1967) realizaram um estudo experimental para determinar as curvas de preferência no plano (n_x de amostras numa direção x , n_q de bits de quantização) e verificaram que essas curvas diferem bastante das curvas de número total de bits constante. De qualquer modo, vale a conclusão importante de que cenas com variação suave de tons de cinza exigem maior número de níveis de quantização mas a amostragem pode ser relativamente espaçada; já no caso de cenas com detalhe fino, a amostragem deve ser densa e a quantização pode ser espaçada.

REFERENCIAS BIBLIOGRÁFICAS

- DAVENPORT, W.B. *Probability and random processes*. Tokyo, Mc Graw Hill, Kogakusha, 1970.
- HECHT, S. The visual discrimination of intensity and the Weber-Fechner Law, *Journal Gen. Physiol.*, vol. 7, 1924, p. 241.
- HUANG, T.S.; TRETIAK, O.J.; PRASADA, B.; YAMAGUCHI, Y. Design considerations in PCM transmission of low resolution monochrome still pictures. *Proceedings of the IEEE*, 55:331-335, 1967.
- JAIN, A.K. Image Data Compression: A Review, *Proceedings of the IEEE*, 69(3):349-389, Mar. 1981.
- KRETZ, F. Subjectively Optimal Quantization of Pictures. *IEEE Transactions Commun.*, COM23(11):1288-1292, Nov. 1975.
- MAX, J. Quantizing for minimum distortion. *IRE Transactions Information Theory*, IT6:7-12, 1960.
- MOIK, J.G. *Digital processing of remotely sensed images*. Washington DC., U.S. Government Printing Office, 1980.
- PANTER, P.F.; DITE, W. Quantizing distortion in pulse-code modulation with nonuniform spacing levels. *Proceedings of the IRE*, 39:44-48, 1951.
- PRATT, W.K. *Digital image processing*. New York, NY, John Wiley, 1978.
- ROBERTS, L.G. Picture coding using pseudo-random noise. *IRE Transactions Information Theory*, IT8(2):145-154, Feb. 1962.
- ROSENFELD, A.; KAK, A.C. *Digital picture processing*, 2ª ed., New York, NY, Academic, 1982.
- SAKRISON, D.J. *Communication theory: transmission of waveforms and digital information*. New York, NY., John Wiley, 1968.

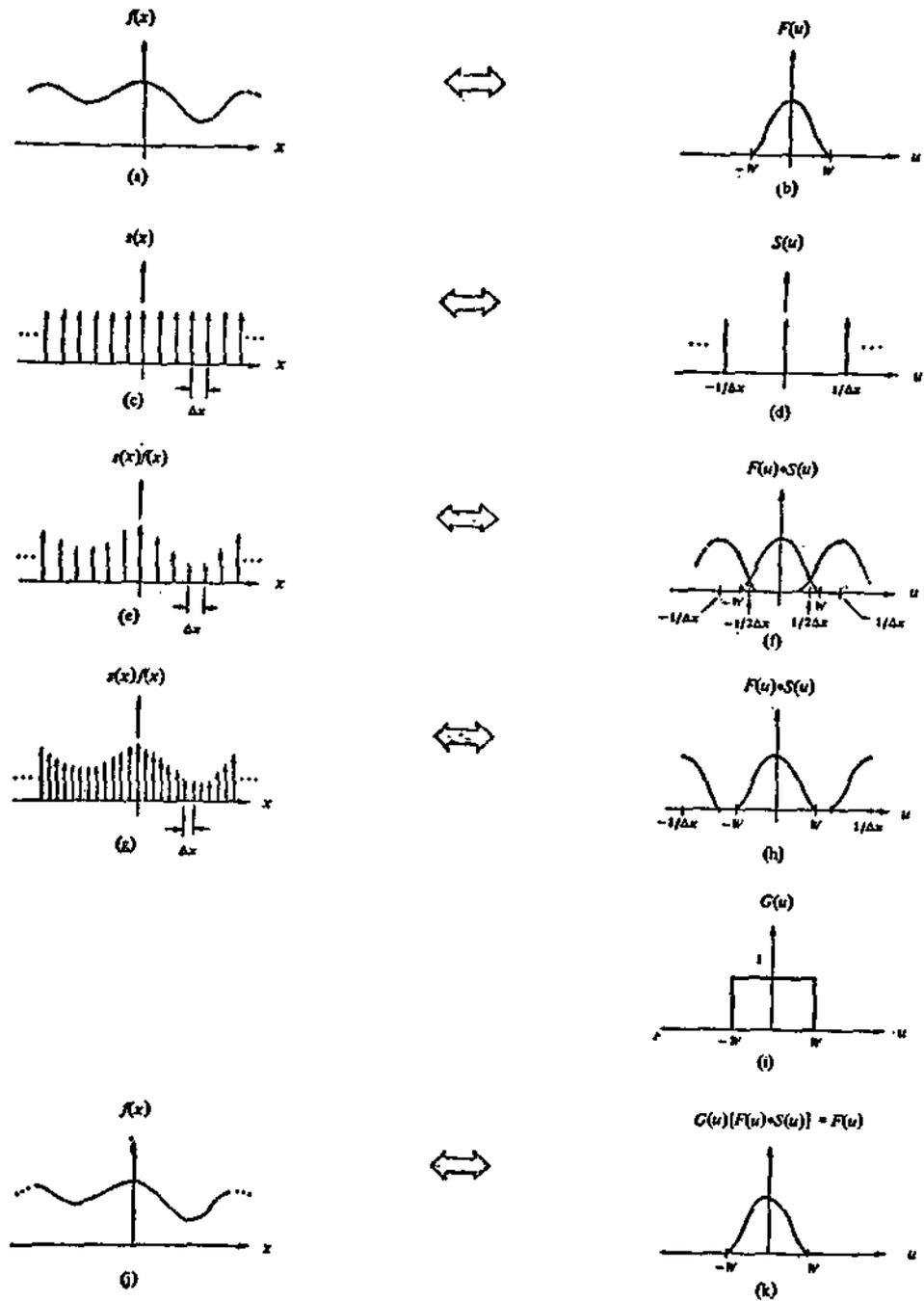


Fig. 2.1 - Ilustração gráfica dos conceitos de amostragem.

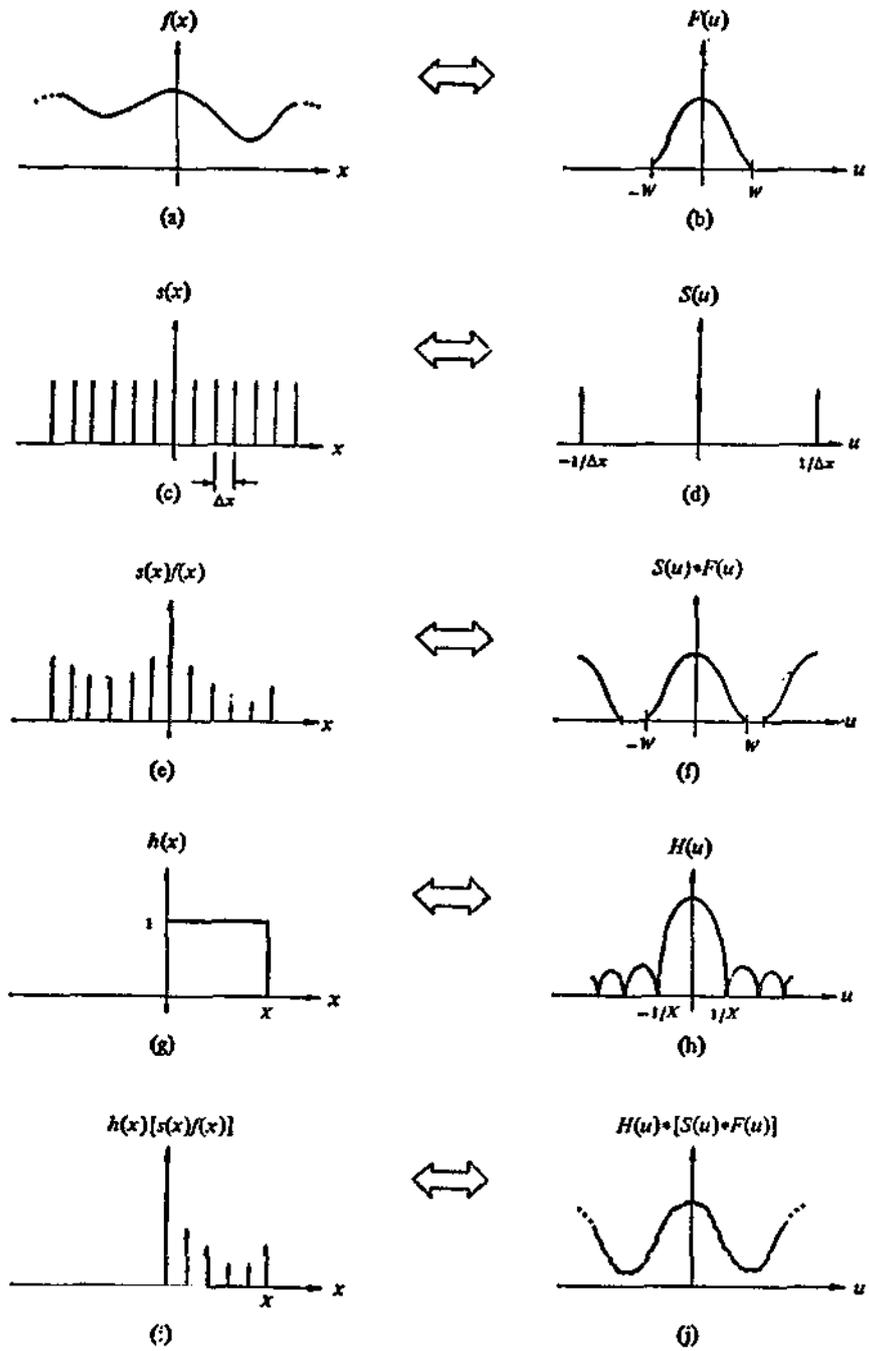


Fig. 2.2 - Ilustração grãfica do efeito de amostragem finita.

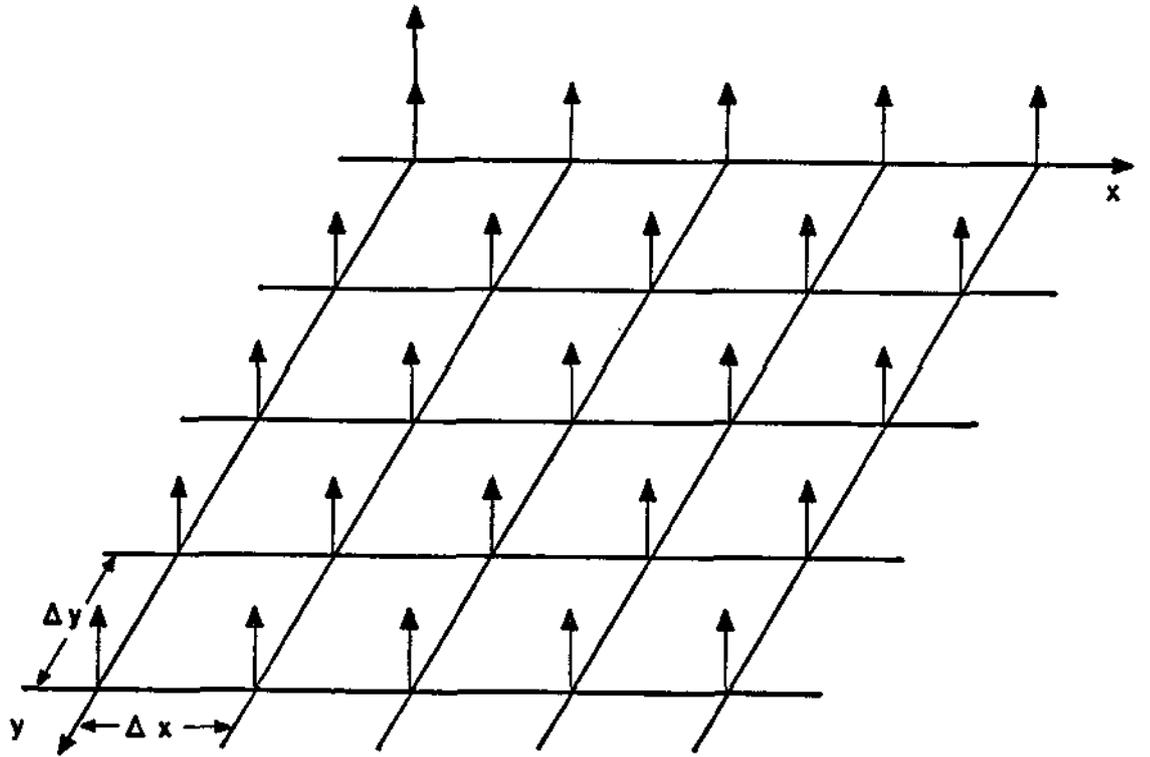


Fig. 2.3 - Matriz bidimensional de amostragem por funções delta.

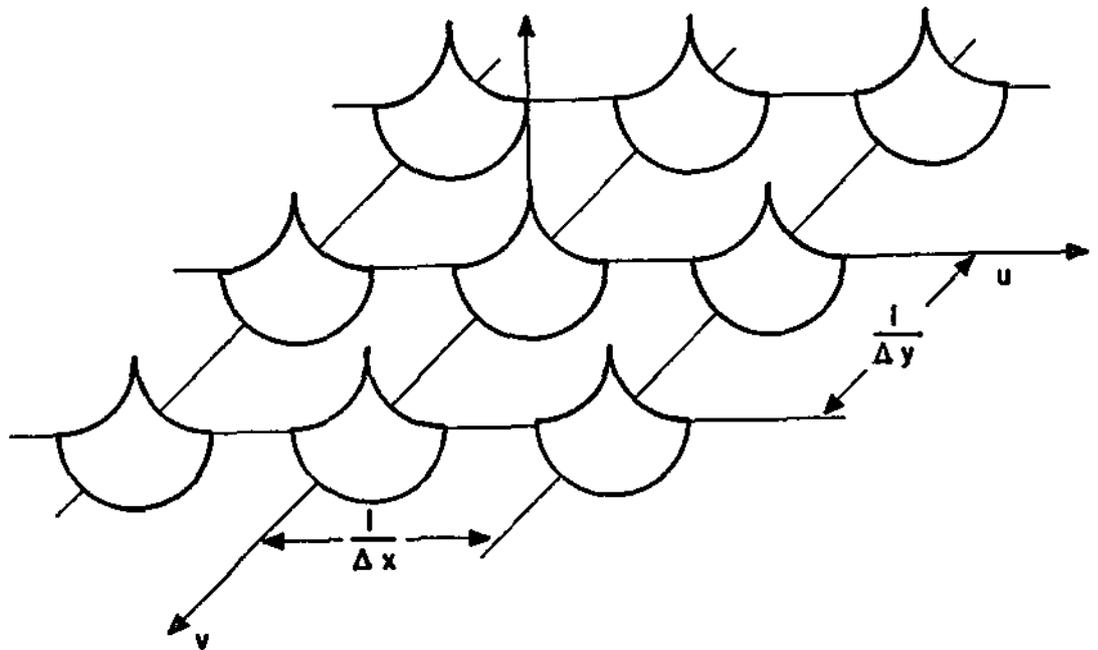
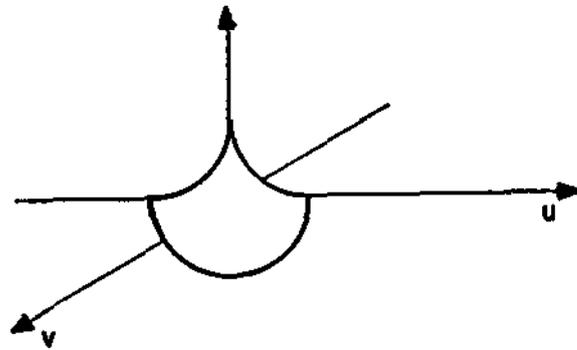
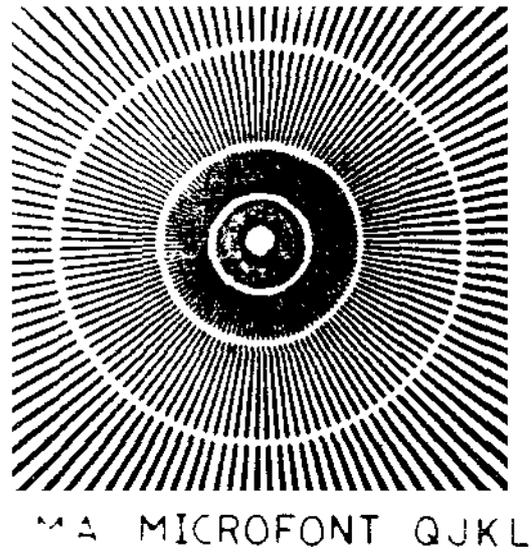
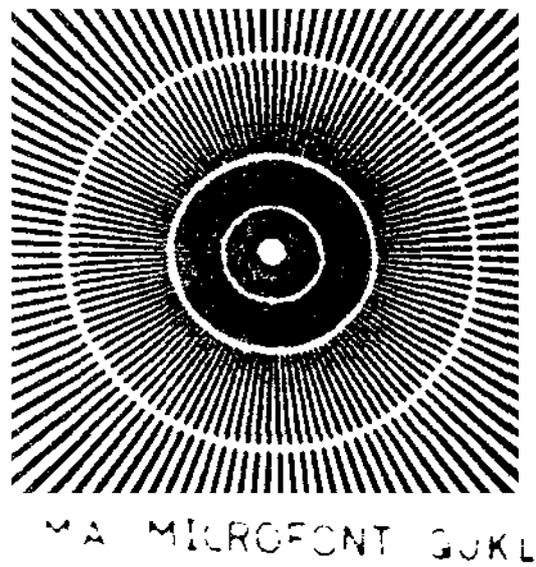


Fig. 2.4 - Espectro de imagem amostrada com largura de faixa limitada: (a) imagem original; (b) imagem amostrada.



(a)



(b)

Fig. 2.5 - Exemplo de padrões de Moiré: (a) imagem original; (b) imagem amostrada.



(a)



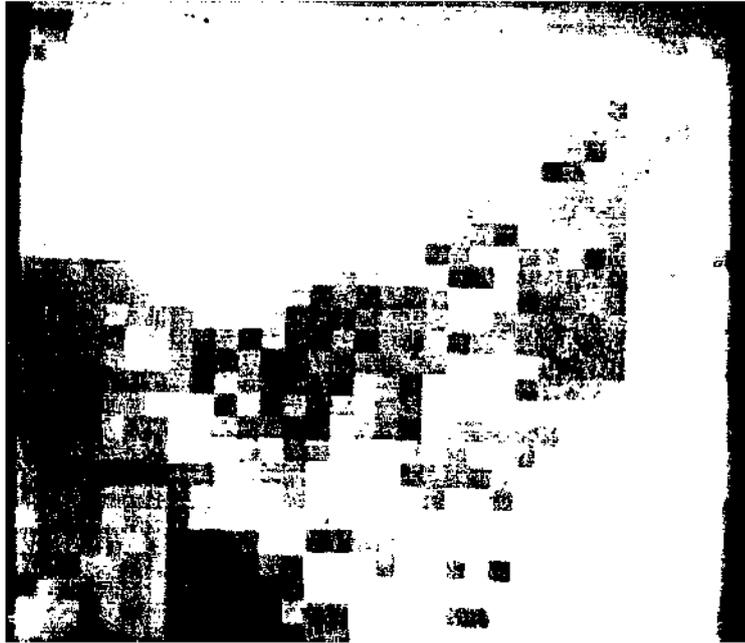
(b)



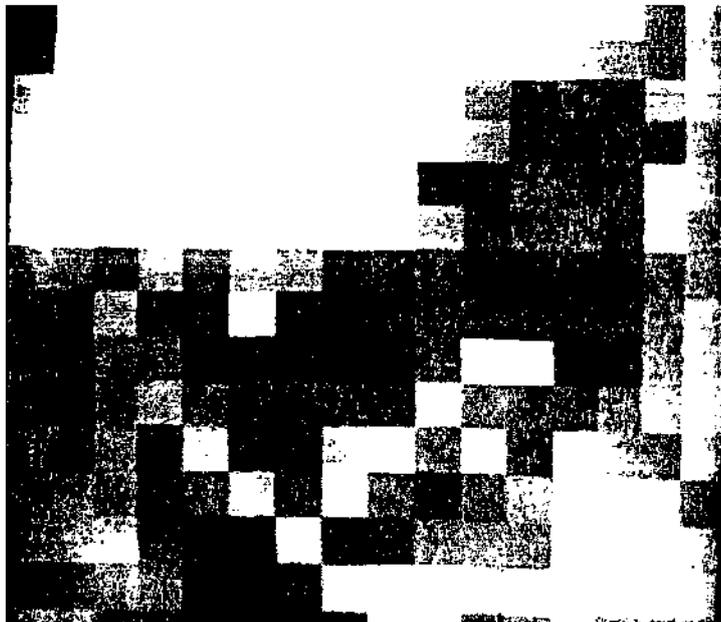
(c)



(d)



(e)



(f)

Fig. 2.6 - Efeitos de redução de número de elementos de imagem: (a) 512 x 512; (b) 256 x 256; (c) 128 x 128; (d) 64 x 64; (e) 32 x 32; (f) 16 x 16.

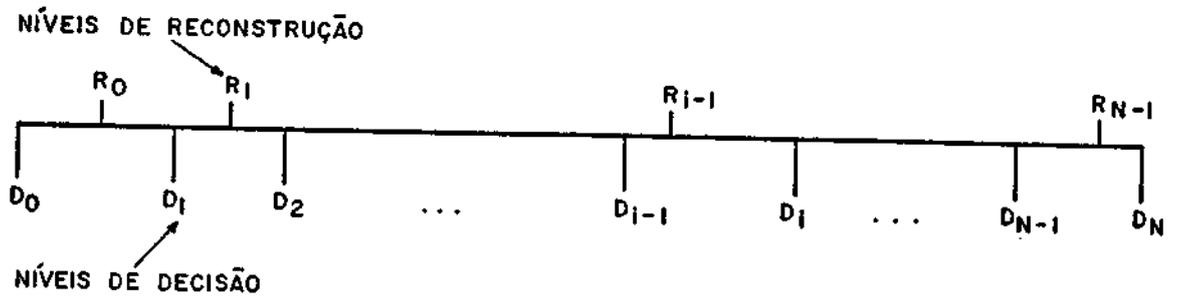


Fig. 2.7 - Representação do processo de quantização de um escalar.

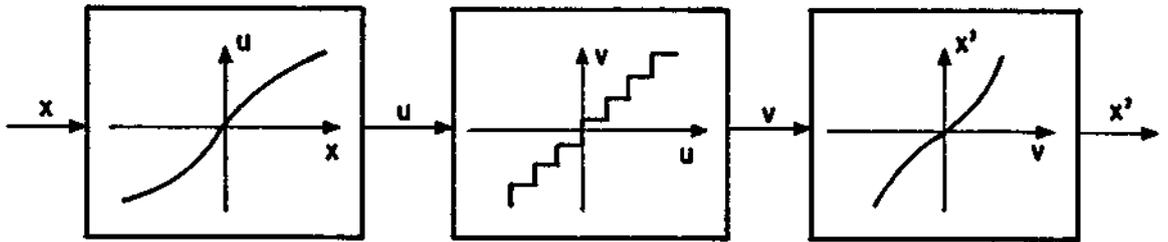


Fig. 2.8 - Processo de compressão-expansão na quantização de um escalar.



(a)



(b)



(c)



(d)

Fig. 2.9 - Quantização não-uniforme: (a) 16 níveis igualmente espaçados; (b) 16 níveis desigualmente espaçados; (c) 4 níveis igualmente espaçados; (d) 4 níveis desigualmente espaçados.



(a)



(b)



(c)



(d)



(e)



(f)

Fig. 2.10 - Efeito visual de redução do número de níveis de quantização: (a) 256 níveis; (b) 32 níveis; (c) 16 níveis; (d) 8 níveis; (e) 4 níveis; (f) 2 níveis.



(a)



(b)

Fig. 2.11 - Efeito de redução de falsos contornos por ruído pseudo-aleatório; (a) PCM com 8 níveis; (b) PCM com 8 níveis e ruído pseudo-aleatório.



A

(a)

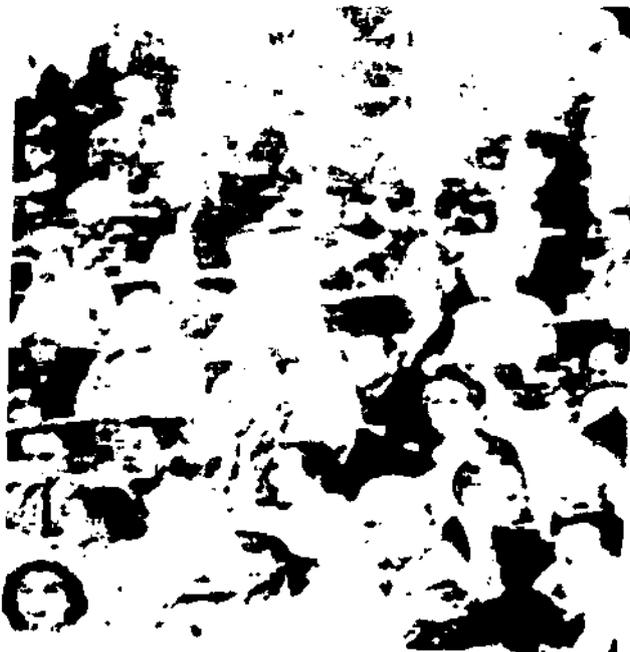


B

(b)



(c)



(d)

Fif. 2.12 - Troca entre amostragem e quantização: (a) imagem de rosto, 128 x 128 amostras, 64 níveis; (b) imagem de rosto, 256 x 256 amostras, 16 níveis; (c) imagem de multidão, 128 x 128, 64 níveis; (d) imagem de multidão, 256 x 256 amostras, 16 níveis.

CAPÍTULO 3

REALCE

Podem-se distinguir dois tipos de transformações radiométricas, ou seja, transformações que modificam o valor dos níveis de cinza dos pontos: restauração - que visa corrigir alguma distorção sofrida pela imagem; realce - que procura enfatizar alguma característica de interesse da imagem. Por vezes estes dois objetivos produzem procedimentos coincidentes. É o caso, por exemplo, quando uma imagem sofreu uma distorção que diminuiu seu contraste, uma transformação que realce as bordas dos objetos das imagens pode, de fato, restaurá-la.

Embora muitas das técnicas de restauração e realce sejam as mesmas (por exemplo, filtragem), os objetivos e enfoques divergem num e noutro caso. O procedimento geral da restauração (Capítulo 4) é a modelagem do processo de distorção para tentar invertê-lo. No realce não há esta preocupação. Como será visto, as técnicas de realce são, na maioria, heurísticas, não havendo mesmo compromisso com a imagem original.

Neste capítulo serão considerados dois tipos de transformações de nível de cinza:

- modificação da escala de cinza (operações pontuais);
- filtragem (operações locais).

Na modificação da escala de cinza, o novo valor do ponto depende só do valor antigo deste ponto, enquanto na filtragem o novo valor depende também dos valores dos outros pontos.

A filtragem pode, por sua vez, ser linear (caso consista em operações lineares na imagem) ou não. Podem ter por objetivo a suavização ("smoothing") ou aumento do contraste ("sharpening") da imagem original.

Embora importantes, não serão estudados os mapeamentos de uma imagem em pseudo-cor e a transformação de componentes principais. No mapeamen

to de pseudo-cor, os n \bar{v} eis de cinza de uma imagem s \bar{a} o mapeados em triplas de n \bar{v} eis de cinza, que correspondem \bar{a} s cores vermelho, verde e azul (outras cores prim \bar{a} rias s \bar{a} o tamb \bar{e} m usadas). A raz \bar{a} o para esta transforma \bar{c} o \bar{e} que o olho humano \bar{e} muito mais sens \bar{i} vel a varia \bar{c} o \bar{e} s de tonalidade que a varia \bar{c} o \bar{e} s de n \bar{v} el de cinza. A transforma \bar{c} o de componentes principais \bar{e} usada em imagens multiespectrais cujas bandas est \bar{a} o correlacionadas, ou seja, cada ponto da imagem \bar{e} uma \bar{e} nupla de n \bar{v} eis de cinza e os valores desta \bar{e} nupla n \bar{a} o s \bar{a} o estatisticamente independentes. A transforma \bar{c} o de componentes principais \bar{e} uma transforma \bar{c} o linear neste espa \bar{c} o de n \bar{v} eis de cinza que produz imagens com bandas n \bar{a} o-correlacionadas e com a informa \bar{c} o concentrada em apenas algumas bandas. Deste modo pode-se trabalhar com imagens com menos bandas mas com o mesmo conte \bar{u} do de informa \bar{c} o.

3.1 - TRANSFORMA \bar{C} AO DA ESCALA DE CINZA

A transforma \bar{c} o radiom \bar{e} trica mais s \bar{i} mplies, e nem por isto a menos \bar{u} til, \bar{e} o mapeamento direto dos n \bar{v} eis de cinza. Num mapeamento direto, os n \bar{v} eis de entrada variam de 0 a N-1 e os de sa \bar{i} da do mapeamento de 0 a M-1 (em geral N = M). A fun \bar{c} o de mapeamento pode ser dada por uma f \bar{o} rmla, por exemplo, uma express \bar{a} o aritm \bar{e} tica, ou atrav \bar{e} s de uma tabela. A Figura 3.1 mostra algumas possibilidades interessantes.

A fun \bar{c} o mostrada na Figura 3.1a real \bar{c} a as varia \bar{c} o \bar{e} s dentro do intervalo [a, b] enquanto "satura" em 0 os valores menores que a e em M-1 os valores maiores que b. Na Figura 3.1b, os n \bar{v} eis de 0 a N-1 s \bar{a} o "comprimidos" para o intervalo [a', b']. Uma limiariza \bar{c} o (ou classifica \bar{c} o) \bar{e} o efeito da fun \bar{c} o mostrada na Figura 3.1c. A imagem resultante tem s \bar{o} dois n \bar{v} eis: 0 e M-1. O efeito da fun \bar{c} o da Figura 3.1d \bar{e} destacar os pontos da imagem no intervalo [a, b]; a imagem resultante tem tamb \bar{e} m dois n \bar{v} eis. A fun \bar{c} o da Figura 3.1e \bar{e} semelhante \bar{a} da Figura 3.1d, mas h \bar{a} uma manuten \bar{c} o dos n \bar{v} eis de cinza fora do intervalo [a, b].

A Figura 3.2 mostra uma fun \bar{c} o que combina v \bar{a} rios aspectos das fun \bar{c} o \bar{e} s mostradas na Figura 3.1. \bar{E} a "curva MB", bastante usada por meteorologistas para real \bar{c} ar os topos das nuvens em imagens (infravermelhas) de sat \bar{e} lites meteorol \bar{o} gicos. A Figura 3.3 mostra o realce de uma imagem pela curva MB.

Embora simples, os mapeamentos diretos são úteis em várias aplicações. Por exemplo, a calibração da curva de resposta de um sensor com a intensidade de energia incidente pode ser feita usando um mapeamento conveniente; o efeito de certas distorções atmosféricas ("espalhamento") em imagens de satélite pode ser minimizado simplesmente subtraindo um valor constante do nível de cinza de todos os pontos da imagem, saturando em 0 os valores negativos ("haze removal"). O realce de imagens que no processo de aquisição tenham sido sub ou sobreexpostas pode ser feito por mapeamentos diretos do tipo mostrado na Figura 3.1a.

3.2 - MODIFICAÇÃO DE HISTOGRAMA

A escolha de mapeamento direto adequada é, em geral, essencialmente empírica. Contudo existe uma classe de métodos - modificação de histograma - nos quais o mapeamento tem por finalidade produzir uma imagem cujo histograma tenha um formato desejado.

Dada uma imagem digital I com K níveis, o histograma de I é uma função $h: [0, K-1] \rightarrow Z_+$ (conjunto dos inteiros não-negativos) onde $h(z)$ dá o número de pontos de I com nível de cinza igual a z .

O histograma de uma imagem contém uma informação global sobre os "objetos" nela contidos. Se todos os pontos da imagem são de um mesmo objeto, o histograma dá uma idéia da probabilidade condicional $p(z/\text{objeto})$ de um ponto possuir um dado nível de cinza z dado que pertence ao objeto. Como em geral espera-se encontrar vários objetos distintos, o histograma reflete uma soma ponderada das probabilidades condicionais. Para um único objeto é comum a distribuição "normal" ("gaussiana"). Imagens constituídas de dois objetos (por exemplo, objeto principal e fundo) podem, com frequência, apresentar histogramas bimodais. Neste caso, uma técnica simples para separar os dois objetos é limiarizar a imagem em algum ponto entre os dois picos.

As Figuras 3.4 e 3.5 mostram imagens com os histogramas respectivos. Na Figura 3.6 a imagem da Figura 3.4 é limiarizada num valor entre os dois picos.

Seja I uma imagem quantizada em K níveis ($[0, K-1]$) com N linhas e M colunas. Neste caso o número total de pontos da imagem é $N.M$. Seja h o histograma de I . $\sum_{z=0}^r h(z)$ dá o número de pontos com nível de cinza menor ou igual a r . Naturalmente $\sum_{z=0}^{K-1} h(z) = N.M$.

O problema da modificação de histograma pode ser posto como o de alterar os valores dos níveis de cinza de I , obtendo-se uma imagem I' tal que h' (histograma de I') tenha o formato desejado. Exposto desta forma, o problema tem uma solução imediata: tomam-se os $h'(0)$ primeiros pontos de I e mapeia-se seu nível de cinza em 0; os $h'(1)$ pontos seguintes são mapeados em 1 e assim por diante. O inconveniente é que a imagem resultante I' muito provavelmente pouco se parecerá com I . Assim, uma restrição adicional à transformação de I em I' é a seguinte: sejam z, z' os níveis de cinza de dois pontos de I , e $f(z)$ e $f(z')$ os níveis de cinza dos mesmos pontos em I' . Então, se $z < z'$, deve-se ter necessariamente $f(z) < f(z')$. Em outras palavras, deve haver uma "monotonicidade" do mapeamento.

Uma restrição adicional é exigir "funcionalidade" para f , ou seja, se $z=z'$, então $f(z)=f(z')$. Neste caso o problema de modificação de histograma pode não ter solução, como poderá ser visto se I tiver dois pontos, ambos com nível de cinza igual a 0, e I' tiver dois pontos, mas com níveis 0 e 1. Para mapeamentos funcionais deve-se procurar obter o histograma que mais se aproxima do histograma desejado.

Para transformar exatamente, de forma monotônica, um histograma noutro, pode-se especificar uma função f com dois argumentos z e z' ($z = 0, \dots, K-1$), onde $f(z, z')$ dá o número de pontos com nível z que são mapeados em z' . As restrições a f são:

$$1) \sum_{z'=0}^{K'-1} f(z, z') = h(z);$$

$$2) \sum_{z=0}^{K-1} f(z, z') = h'(z);$$

$$3) z_1 < z_2 \text{ e } f(z_1, z') > 0 \text{ e } z'' < z' \text{ então } f(z_2, z'') = 0.$$

A restrição 1 implica que todos os pontos (e não mais) de um certo nível z são mapeados em algum nível z' . A restrição 2 assegura que o histograma obtido tem o formato desejado ($h'(z')$). Finalmente, a restrição 3 implica à monotonicidade do mapeamento, pois senão, sejam $z_1 < z_2$, $f(z_1, z_2) > 0$ e $z'' < z'$ e $f(z_2, z'') > 0$, e p_1 e p_2 dois pontos de I cujos níveis de cinza sejam z_1 e z_2 mas que sejam mapeados em z' e z'' , respectivamente (o que é possível, pois $f(z_1, z')$ e $f(z_2, z'') > 0$). Neste caso tem-se $z'' < z'$, o que contraria a hipótese.

Exemplo 3.1

Seja uma imagem cujo histograma é dado por:

z	I	0	I	1	I	2	I	3	I	4	I	5	I	6	I	7	I
$h(z)$	I	1	I	7	I	21	I	35	I	30	I	18	I	12	I	4	I

e deseja-se obter uma imagem I' cujo histograma é plano, isto é, $h'(z') = 16$, $z' = 0, 1, \dots, 7$. A função $f(z, z')$, $z = 0, 1, \dots, 7$ que realiza este mapeamento está dada na Tabela 3.1 (as entradas não preenchidas correspondem a $f(z, z') = 0$).

O algoritmo 3.1 constrói a função $f(z, z')$, dados $h(z)$ e $h'(z')$.

Algoritmo 3.1 - Modificação de histograma.

1) Para todo z, z' faça $f(z, z') \leftarrow 0$;

2) $z' \leftarrow 0$;

3) $H' \leftarrow h'(z')$;

4) Para $z \leftarrow 0$ até $K-1$ faça:

4.1) $H \leftarrow h(z)$;

4.2) Enquanto $H > 0$ faça:

4.2.1) $f(z, z') \leftarrow \min(H, H')$;

4.2.2) $H \leftarrow H - f(z, z')$;

4.2.3) $H' \leftarrow H' - f(z, z')$;

4.2.4) Se $h' \neq 0$, faça:

4.2.4.1) $z' \leftarrow z' + 1$;

4.2.4.2) $H' \leftarrow h'(z')$

TABELA 3.1

FUNÇÃO DE MAPEAMENTO DE NÍVEIS DE CINZA

z'	0	1	2	3	4	5	6	7	
z									
0	1								1
1	7								7
2	8	13							21
3		3	16	16					35
4					16	14			30
5						2	16		18
6								12	12
7								4	4
	16	16	16	16	16	16	16	16	16

O mapeamento da imagem propriamente dito pode ser feito percorrendo a imagem original I e, dado um ponto com nível de cinza z , mapeando-o em z' desde que $f(z, z') > 0$; em cada mapeamento subtrai-se 1 de $f(z, z')$.

O Algoritmo 3.1 tem o inconveniente de produzir mapeamentos não-funcionais, isto é, um nível de cinza pode ser mapeado em vários níveis da imagem transformada. Isto cria o problema de, dado um ponto com um certo nível de cinza, saber qual nível mapeá-lo de fato. Pode-se pensar em algumas soluções para este problema. A primeira é, no percorrimento da imagem original, se $f(z, z'-1) = 0$ e $f(z, z') > 0$, mapear os primeiros $f(z, z')$ pontos de nível de cinza em z' , os seguintes $f(z, z'+1)$ em $z'+1$ e assim sucessivamente. Outro modo é escolher aleatoriamente o nível z' (entre os que tenham $f(z, z') > 0$) para um dado ponto p com nível de cinza z . Em geral, contudo, não se nota o efeito da não-funcionalidade do mapeamento e ambas as soluções conduzem a resultados visualmente equivalentes.

Uma importante aplicação do Algoritmo 3.1 é a produção de imagens "equalizadas", isto é, imagens cujo histograma é plano. Como pode ser visto na Figura 3.7, a equalização pode ter um efeito notável de realce da imagem. No caso da Figura 3.7, isto é devido a que, na imagem original, os níveis de cinza ocupam um intervalo pequeno da escala total de níveis de cinza, enquanto na imagem equalizada toda escala é utilizada.

3.3 - FILTRAGEM

Na filtragem, o nível de cinza de um ponto p , após a transformação, depende do valor do nível de cinza original do ponto e de outros pontos da vizinhança de p . Esta vizinhança não é necessariamente limitada, mas em geral os pontos mais próximos de p contribuem mais para o novo nível de pontos mais afastados.

Dadas duas imagens $I(x, y)$ e $J(x, y)$, $x=0, 1, \dots, M-1$ e $y=0, 1, \dots, N-1$, a "convolução" (discreta) destas duas imagens é uma outra imagem $H(x, y)$, $x=0, 1, \dots, M-1$ e $y=0, 1, \dots, N-1$ dada por:

$$H(x, y) = (1/M.N) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \cdot J(x-m, y-n), \quad (3.1)$$

ou equivalentemente por:

$$H(x, y) = (1/M.N) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(x-m, y-n) \cdot J(m, n), \quad (3.2)$$

onde os valores fora dos intervalos $[0, M-1]$ e $[0, N-1]$ são dados pelas fórmulas:

$$\begin{aligned} F(x, -y) &= F(x, N-y), \\ F(-x, y) &= F(M-x, y), \end{aligned} \tag{3.3}$$

onde F pode ser I ou J .

As fórmulas 3.3 são equivalentes a considerar uma imagem $F(x, y)$ definida em $[0, M-1]$ e $[0, N-1]$ como definida em $[-M, M-1]$ e $[-N, N-1]$. De um modo geral, pode-se pensar que I , J e H são definidas para todo x e y inteiros e que as três imagens se repetem em intervalos de M pontos na vertical e de N pontos na horizontal.

O resultado da convolução de I por J num ponto p é na realidade de uma média ponderada dos pontos de I , onde os pesos são dados pela imagem J (veja Equação 3.2). A Figura 3.8 mostra, dado um ponto p de I , quais são os pontos correspondentes que servem de peso numa vizinhança 3×3 . A importância da convolução deriva do fato que toda transformação linear invariante a deslocamento (veja Capítulo 1) é equivalente à convolução desta imagem com uma outra. Dadas duas imagens I e J , usa-se representar H , convolução de I por J , por $I * J$.

A transformada (discreta) de Fourier de uma imagem $I(x, y)$ (veja Apêndice C) é uma outra imagem $I(u, v)$ (notação: $F(I(x, y))$) cujos valores são complexos, definidos numa mesma região, obtida pela fórmula:

$$I(u, v) = (1/M.N) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) \exp[-j.2.\pi. ((m.u)/M + (n.v)/N)] \tag{3.4}$$

A transformada de Fourier está relacionada com a convolução, uma vez que a transformada de Fourier da convolução de duas imagens I e J é igual ao produto das transformadas de Fourier de I e J , como indicado na Equação 3.5.

$$F(I \star J) = F(I) \cdot F(J) \quad (3.5)$$

Deve-se observar que o produto mostrado na Equação 3.5 é o produto ponto a ponto e não o produto matricial.

As operações de filtragem podem ser divididas em duas classes: a filtragem linear, obtida através uma operação de convolução, e a filtragem não-linear. Em ambos casos, como já foi mencionado, o valor final de um ponto depende dos valores de seus pontos vizinhos, bem como da relação de vizinhança existente. Há também a filtragem recursiva (tal como o filtro de Kalman, por exemplo) que é linear mas não invariante a deslocamento. Como este tipo de filtragem é usado principalmente na restauração, ele será tratado no Capítulo 4.

3.3.1 - Filtragem por convolução

O efeito da convolução de uma imagem I por J, é, em geral, mais bem compreendido estudando o comportamento do produto das transformadas de Fourier de I e J. Tal como na convolução, deve-se estender o domínio de definição da imagem e de sua transformada para todo o plano da forma: $I(x+m.M, y+n.N) = I(x, y)$ e $I(u+m.M, v+n.N) = I(u, v)$. A fórmula para o câlculo da transformada, neste caso, é a mesma (Equação 3.5).

Caso a imagem I tenha valores reais, há uma repetição nos valores da transformada de I. Na realidade, pode-se mostrar que:

$$I(u, v) = I^* (-u+m.M, -v+n.N),$$

onde I^* é o conjugado de I.

Desta relação segue que, para $u=0$, $I(0, v) = I^*(0, N-v)$; para $v=0$, tem-se que $I(u, 0) = I^*(M-u, 0)$; e para $u \neq 0$ e $v \neq 0$, $I(u, v) = I^*(M-u, N-v)$. A Figura 3.9 mostra para $M=5$ e $N=6$ quais os termos conjugados na transformada de Fourier de uma imagem real.

Por analogia com a transformada contínua de Fourier; os índices u e v são chamados "frequências espaciais". Pode-se ver, da fórmula da antitransformada,

$$I(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} I(u, v) \cdot \exp[j \cdot 2 \cdot \pi \cdot ((x \cdot u)/M + (y \cdot v)/N)], \quad (3.6)$$

que $I(u, v)$ é o coeficiente de

$$B(x, y; u, v) = \exp[j \cdot 2 \cdot \pi \cdot ((x \cdot u)/M + (y \cdot v)/N)],$$

que pode ser decomposto em suas componentes cossenoidais e senoidais:

$$\begin{aligned} B(x, y; u, v) &= \cos[2 \cdot \pi \cdot ((x \cdot u)/M + (y \cdot v)/N)] + j \cdot \text{sen}[2 \cdot \pi \cdot ((x \cdot u)/M + (y \cdot v)/N)] \\ &= BC(x, y; u, v) + j \cdot BS(x, y; u, v). \end{aligned}$$

Para um par (u_0, v_0) fixo, BC e BS são funções bidimensionais. É interessante observar o comportamento das funções BCs e BSs. Fixando uma coluna, por exemplo, $y=0$, tem-se que:

$$BC(x, u) = \cos [(2 \cdot \pi \cdot x \cdot u)/M] \quad \text{e}$$

$$BS(x, u) = \text{sen} [(2 \cdot \pi \cdot x \cdot u)/M].$$

A Figura 3.10 mostra para $M=8$ o gráfico de BC com x e u variando de 0 a 7. A variação de BS é semelhante, havendo só uma alteração de fase.

Pode-se ver na Figura 3.10 que só para valores de u próximos de 0 e M as funções-base são aproximações (grosseiras) de funções senoidais. Também, para valores de u próximos de 0 e M , a variação da função-base com x é mais lenta do que para valores próximos de $M/2$. Em outras palavras, valores de (u, v) próximos de $(0, 0)$, $(0, N)$, $(M, 0)$ e (M, N) correspondem a "baixas frequências", e valores de (u, v) distantes destes pontos correspondem a "altas frequências". As partes hachuradas na Figura 3.11 mostram as re

ções de baixas frequências numa transformada de Fourier. Note também que a maior frequência contida nas funções-base mostradas na Figura 3.10 é de $1/2$ ciclo por ponto, que corresponde a $u=4$.

Como a transformada de Fourier é em geral complexa, não é possível visualizá-la diretamente como uma imagem. Assim, para efeito de apresentação, costuma-se construir a partir da transformada uma imagem cujos valores dos pontos são o módulo da transformada. Além disto, prefere-se usar não $I(x, y)$, mas $I(x, y) \cdot (-1)^{x+y}$. Isto faz com que a transformada de Fourier fique deslocada de $(M/2, N/2)$; portanto, a região de baixa frequência passa a ser o meio da imagem, o que facilita a visualização.

Como foi dito, a convolução de uma imagem $I(x, y)$ com outra imagem (filtro) $J(x, y)$ pode ser computada tomando as transformadas de Fourier $F(I)$ e $F(J)$, multiplicando-as e achando a antitransformada do produto. Assim, conhecendo a transformada de Fourier de um filtro $J(x, y)$, pode-se prever o efeito deste quando aplicado a uma imagem qualquer.

É comum dividir os filtros em duas grandes classes:

1) passa-baixas,

2) passa-altas,

dependendo de quais frequências são atenuadas pelo uso do filtro. Um filtro passa-baixas tem valores próximos de zero para as altas frequências; um filtro passa-altas tem valores próximos de zero para as baixas frequências. Há ainda um terceiro tipo - passa-faixa - que atenua tanto as frequências altas quanto as baixas, preservando contudo as frequências médias.

O efeito visual de um filtro passa-baixas é o de "suavização" ("smoothing") da imagem, uma vez que as altas frequências, que correspondem às transições abruptas, são atenuadas. A suavização tende também, pelas mesmas razões, a minimizar o efeito do ruído em imagens. A filtragem passa-baixa tem, por outro lado, o efeito indesejado de diminuir a resolução da imagem.

Para filtros passa-altas, o efeito obtido é em geral o de "agudização" ("sharpening") da imagem: as transições entre regiões diferentes tornam-se mais nítidas. O efeito indesejado é o de enfatizar o ruído porventura existente na imagem.

Embora se possa efetuar a convolução de duas imagens usando as suas transformadas de Fourier, para filtros que são nulos - a não ser em uma pequena região ("núcleo" ou "máscaras") - é preferível fazer a operação "no domínio do espaço", ou seja, com as imagens originais. Seja, por exemplo, a operação de "média não-ponderada" numa vizinhança 3x3; esta operação corresponde à convolução com a imagem da Figura 3.12. O núcleo do filtro média é:

$$\begin{array}{ccc} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9 \end{array}$$

Fig. 3.12 - Imagem do filtro da média.

Na operação feita no domínio do espaço devem-se calcular os pontos pertencentes à borda de modo diferente dos demais, já que estes não dispõem de todos os vizinhos. Para ter um resultado equivalente ao da transformada de Fourier, deve-se considerar a imagem como periódica, o que não é muito realista. É preferível, pois, dar um tratamento diferenciado aos pontos de borda.

Por questões de simetria usam-se, na definição dos núcleos dos filtros, vizinhanças $n \times n$, onde n é um número ímpar. Por questões de eficiência computacional, preferem-se valores pequenos para n (no máximo 7). É fácil de ver que, se a soma dos pesos de um núcleo for igual a 1, a média da imagem resultante será igual à média da imagem original. De um modo geral, pode-se dizer que a média da imagem filtrada será igual à da imagem original multiplicada pelo somatório dos pesos do núcleo. Seja m' a média de uma imagem filtrada $I'(x, y)$; então:

$$m' = (1/M.N) \sum_x \sum_y I'(x, y).$$

Seja também $V(x, y)$ a vizinhança de (x, y) que contribui para o ponto (x, y) , $V_i(x, y)$ o $i^{\text{ésimo}}$ vizinho de (x, y) , e p_i o peso associado a este vizinho. Neste caso tem-se:

$$m' = (1/(M.N)) \sum_x \sum_y \left[\sum_{V(x, y)} p_i \cdot V_i(x, y) \right]$$

À exceção dos pontos próximos às bordas, todo ponto de I é multiplicado uma vez por cada peso p_i . Supondo que a imagem se repete ciclicamente, isto é também verdade para os pontos da borda. Pode-se, pois, escrever:

$$\begin{aligned} m' &= (1/(M.N)) \sum_x \sum_y I(x, y) \left[\sum_{V(x, y)} p_i \right] \\ &= \left[\sum_{V(x, y)} p_i \right] (1/(M.N)) \sum_x \sum_y I(x, y) \\ &= \left[\sum_{V(x, y)} p_i \right] \cdot m \\ &= m \end{aligned}$$

onde m é a média da imagem $I(x, y)$, ou seja, antes da filtragem.

Este resultado pode ser mostrado também através da transformada de Fourier, usando a propriedade:

$$I(0, 0) = \sum_x \sum_y I(x, y).$$

Filtros do tipo média (onde todos os pesos são positivos) têm o efeito de suavizar a imagem. A Figura 3.13 mostra a transformada de Fourier para o filtro:

$$\begin{array}{ccc} 0 & 0,20 & 0 \\ 0,20 & 0,20 & 0,20 \\ 0 & 0,20 & 0 \end{array}$$

quando aplicado a uma imagem 6x6. É mostrada apenas a parte real, já que a parte imaginária é nula.

Outros filtros, nos quais alguns pesos são negativos, tendem a realçar diferenças e, portanto, a tornar a imagem aguda. A Figura 3.14 mostra a transformada do filtro:

$$\begin{array}{ccc} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{array}$$

A parte imaginária, como no caso do filtro média, é nula.

Em geral, se o filtro for completamente simétrico em relação ao ponto central do núcleo, a parte imaginária será nula. Isto é fácil de ver, pois a parte imaginária é dada por:

$$I_i(u, v) = \sum_x \sum_y I(x, y) \operatorname{sen}[2.\pi((x.u)/M+(y.v)/N)].$$

Se $u = v = 0$, tem-se $I_i(u, v) = 0$; nos outros casos tem-se:

$$\operatorname{sen}[2.\pi(((M-x).u)/M+((N-y).v)/N)] =$$

$$\operatorname{sen}[2.\pi.(u+v) - 2.\pi((x.u)/M + (y.v)/N)] =$$

$$- \operatorname{sen}[2.\pi((x.u)/M + (y.v)/N)].$$

Se $I(x, y) = I(M-x, N-y)$, ou seja, se a imagem for simétrica, então $I_i(u, v) = 0$.

Realce de características específicas

Filtros de convolução podem ser usados para realçar (ou detectar) certas características presentes na imagem, tais como bordas, linhas, curvas ou manchas ("spots").

A Figura 3.15 mostra máscaras (núcleos) de filtros para a detecção de bordas em vários sentidos.

Norte	1	1	1	Nordeste	1	1	1
	1	-2	1		-1	-2	1
	-1	-1	-1		-1	-1	1
Leste	-1	1	1	Sudeste	-1	-1	1
	-1	-2	1		-1	-2	1
	-1	1	1		1	1	1
Sul	-1	-1	-1	Sudoeste	1	-1	-1
	1	-2	1		1	-2	-1
	1	1	1		1	1	1
Oeste	1	1	-1	Noroeste	1	1	1
	1	-2	-1		1	-2	-1
	1	1	-1		1	-1	-1

Fig. 3.15 - Máscaras de filtros direcionais de bordas.

O realce de bordas, independentes da direção, pode ser obtido através das máscaras dadas na Figura 3.16.

0	-1	0	-1	-1	-1	1	-2	1
-1	4	-1	-1	8	-1	-2	4	-2
0	4	0	-1	-1	-1	1	-2	1
	a			b			c	

Fig. 3.16 - Filtros não-direcionais de bordas.

As máscaras dadas na Figura 3.16 são aproximações do operador laplaciano aplicado à função bidimensional que é a imagem. Se as derivadas direcionais são aproximadas pelas diferenças

$$\text{DELTA}_X I(x, y) = I(x-1, y) - I(x, y),$$

$$\text{DELTA}_Y I(x, y) = I(x, y-1) - I(x, y),$$

as derivadas ficam:

$$\begin{aligned} \text{DELTA}_X^2 I(x, y) &= \text{DELTA}_X I(x+1, y) - \text{DELTA}_X I(x, y) = \\ &= 2 \cdot I(x, y) - I(x-1, y) - I(x+1, y) \end{aligned}$$

e

$$\text{DELTA}_Y^2 I(x, y) = 2 \cdot I(x, y) - I(x, y-1) - I(x, y+1).$$

Pela definição de laplaciano,

$$\text{DELTA}^2 f = \text{DELTA}_X^2 f + \text{DELTA}_Y^2 f,$$

tem-se

$$\text{DELTA}^2 I(x, y) = 4 \cdot I(x, y) - [I(x-1, y) + I(x+1, y) + I(x, y-1) + I(x, y+1)],$$

que corresponde ao filtro da Figura 3.16a.

Linhas em imagens podem ser detetadas através dos filtros de convolução dados pelas máscaras da Figura 3.17. A Figura 3.17a mostra um filtro 3x3 para linhas verticais, enquanto o filtro da Figura 3.17b realça linhas horizontais.

-0,5	↑	-0,5	-0,5	-0,5	-0,5
-0,5	↑	-0,5	1	↑	↑
-0,5	↑	-0,5	-0,5	-0,5	-0,5
	a			b	

Fig. 3.17 - Máscaras para realce de linhas.

O filtro da Figura 3.17 deve ser usado com cautela pois ele pode responder a características que não são linhas, tais como bordas e mesmo picos. Uma solução deste problema é usar operadores não-lineares, o que será visto na Seção 3.3.2.

De modo análogo ao mostrado na Figura 3.17, podem ser construídas máscaras para linhas inclinadas de 45 graus com a horizontal.

3.3.2 - Filtragem não-linear

Sejam os sinais unidimensionais dos tipos "ruído" e "borda" (degrau) mostrados na Figura 3.18. Os efeitos de um filtro de convolução dos tipos (média) $[1/3, 1/3, 1/3]$ e diferença $[-1, 3, -1]$ são mostrados na Figura 3.19. Podem-se ver nestes casos os efeitos de atenuação do ruído (Figura 3.19a), "borramento" da borda (Figura 3.19b), ampliação do ruído (Figura 3.19c) e realce da borda (Figura 3.19d).

Os efeitos indesejados mostrados na Figura 3.19c e 3.19d podem ser minorados usando filtros não-lineares. Nesta seção serão estudados os filtros de mediana, vizinhança seletiva e da moda, os quais suavizam a imagem sem contudo diminuir sua resolução.

No filtro da mediana, os pontos da vizinhança de (x, y) são ordenados, e tomado como novo valor para (x, y) o valor mediano desta ordena

ção. O efeito do filtro da mediana (unidimensional) que envolve e elementos (o ponto e seus dois vizinhos mais próximos) é mostrado na Figura 3.20. Note-se que na Figura 3.20a o ruído foi eliminado (não somente atenuado), enquanto na Figura 3.20b o degrau permaneceu inalterado.

É possível também, em vez de tomar a mediana da vizinhança, escolher o valor máximo ou o valor de ordem qualquer. Caso a imagem seja binária, escolher o valor mínimo pode ser feito com o "e" lógico dos valores da vizinhança, e escolher o valor máximo corresponde ao "ou" lógico.

Uma alternativa que produz resultados interessantes é tomar o valor mais frequente de uma vizinhança - a "moda". No caso dos sinais da Figura 3.18, a aplicação deste filtro dá resultados semelhantes à aplicação do filtro da mediana (Figura 3.20).

Na vizinhança seletiva, procura-se tomar a média não em toda a vizinhança, mas sim numa sub-vizinhança que satisfaça alguma propriedade dada, tal como uniformidade, ou apresente um valor médio próximo do valor do ponto (x, y) . O uso do filtro da vizinhança seletiva presume que a imagem seja composta de regiões razoavelmente uniformes; a seleção da vizinhança procura fazer com que o ponto (x, y) seja operado só com pontos de sua região.

A detecção de características tais como bordas, linhas, curvas e manchas pode ser feita também com filtros não-lineares. Para o realce de bordas, o método não-linear mais simples talvez seja o operador gradiente de Roberts, dado por:

$$GR(x, y) = [I(x, y) - I(x-1)]^2 + [I(x-1) - I(x-1, y)]^2 \quad 1/2.$$

Devido ao custo computacional, as operações de elevar ao quadrado e raiz quadrada são muitas vezes substituídas pelo valor absoluto das diferenças cruzadas:

$$GA(x, y) = [I(x, y) - I(x-1)] + [I(x, y-1) - I(x-1, y)],$$

que é mais eficiente. $GR(x, y)$ está relacionado com $GA(x, y)$ pela desigualdade.

$$GR(x, y) \leq GA(x, y) \leq \sqrt{2} GR(x, y).$$

Uma desvantagem do operador de Roberts é sua "anisotropia", ou seja, sua assimetria. Dependendo da direção, certas bordas são mais realçadas que outras, mesmo tendo igual magnitude.

Um operador gradiente mais sofisticado (3x3) é o operador de Sobel, o qual consiste em:

$$GS(x, y) = (X^2 + Y^2)^{1/2}$$

onde

$$X = (I(x-1, y+1) + 2 \cdot I(x, y+1) + I(x+1, y+1)) - \\ (I(x-1, y-1) + 2 \cdot I(x, y-1) + I(x+1, y-1));$$

$$Y = (I(x-1, y-1) + 2 \cdot I(x-1, y) + I(x-1, y+1)) - \\ (I(x+1, y-1) + 2 \cdot I(x+1, y) + I(x+1, y+1)).$$

O princípio semelhante ao da vizinhança seletiva pode ser usado para realce de características. Seja o operador definido por:

$$LV(x, y) = I(x, y+1) \cdot I(x, y) + I(x, y-1) - (1/2) \cdot (I(x-1, y+1) + I(x+1, y+1) + \\ + I(x-1, y) + I(x+1, y) + I(x-1, y-1) + I(x+1, y-1))$$

se

$$I(x, y+1) > I(x-1, y+1) \\ I(x, y+1) > I(x+1, y+1) \\ I(x, y) > I(x+1, y) \\ I(x, y) > I(x+1, y) \\ I(x, y-1) > I(x-1, y-1) \\ I(x, y-1) > I(x+1, y-1)$$

e por

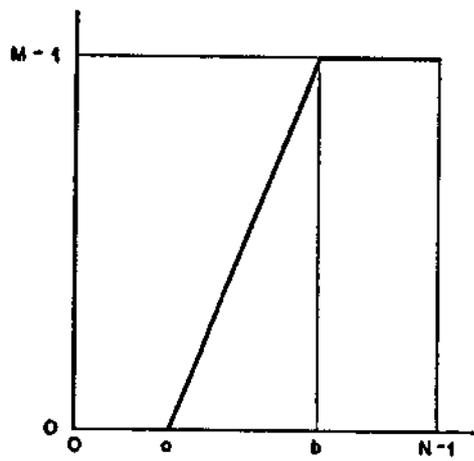
$LV(x, y) = 0$, em caso contrário. Este operador tem a propriedade de realçar linhas verticais mais escuras que o fundo, sem realçar pontos isolados (como os filtros lineares).

Uma análise mais atenta da definição deste operador revela que ele é a implementação do conceito de uma linha vertical escura numa vizinhança 3×3 . O mesmo princípio pode ser usado para implementação de detetores de linhas claras (verticais ou horizontais), ou mesmo de características diferentes como bordas e manchas. Contudo, quanto mais complexa for a característica, mais elaborado deverá ser o filtro que a implementa, o que faz com que torne ineficiente sua implementação em computador.

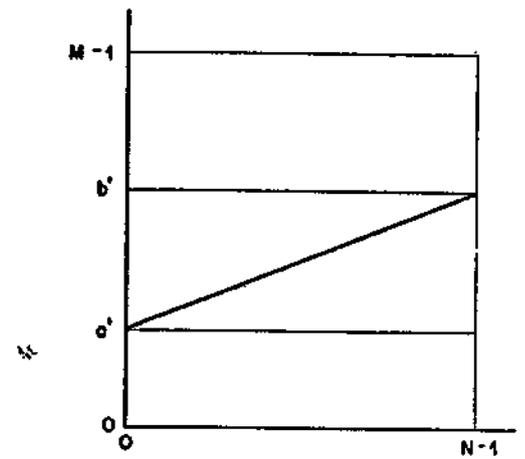
REFERÊNCIAS BIBLIOGRÁFICAS

- ABDOU, I.E.; PRATT, W.K. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of the IEEE*, 67(5):753-763, 1979.
- CHITTINENI, C.B. Edge and line detection in multidimensional noisy imagery data. *IEEE Transactions on Geoscience and Remote Sensing*, GE-21 (4):163-174, 1983.
- DAVIS, L.S.; ROSENFELD, A. Iterative histogram modification. *IEEE Transactions Systems, Man and Cybernetics*, SMC-8:300-302, 1978.
- Noise cleaning by iterated local averaging. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-8:705-710, 1978.
- HALL, E.L. Almost uniform distribution for computer image enhancement. *IEEE Transactions Computers*, C-23(2):207-208, 1974.
- HALL, E.L.; KRUGER, R.; DWYER III, S.J.; HALL, D.L.; MCLAREN, R.W.; LODWICK, G.S. A survey of preprocessing and feature extraction techniques for radiographic images. *IEEE Transactions on Computers*, C-2(9):1032-1044, 1971.
- HUECKEL, M. An operator which locates edges in digital pictures. *Journal of the ACM*, 18:113-125, 1971.
- HUMMEL, R. Histogram modification techniques. *Computer Graphics and Image processing*, 4:209-224, 1975.
- Image enhancement by histogram transformation. *Computer Graphics and processing*, 6(2):184-195, 1977.
- KIRSH, R. Computer determination of the constituent Structure of biological images. *Computers and Biomedical Research*, 4(3):315-328, 1971.
- MASCARENHAS, N.D.A.; PRADO, L.O.C. A Bayesian approach to edge detection in images. *IEEE Transactions Automatic Control*, AC-25(2):36-43, 1980.
- MODESTINO, J.W.; FRIES, R.W. Edge detection in noisy images using recursive digital filtering. *Computer Graphics and Image Processing*, 6(5):409-433, 1977.

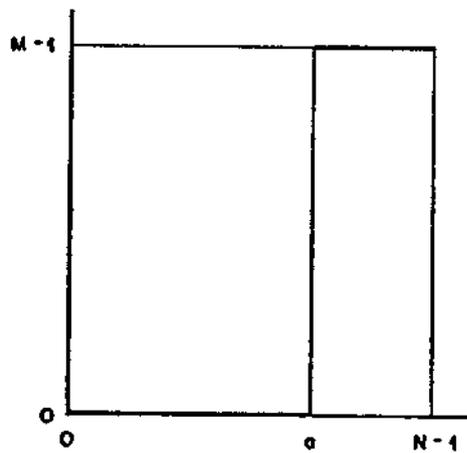
- NAHI, N.E. Role of recursive estimation in statistical image enhancement. *Proceeding of the IEEE*, 60(7):872-877, 1972.
- NATHAN, R. Spatial frequency filtering. In: Lipkin, B.S.; Rosenfeld, A. *Picture Processing and Psychopictorics*. New York, Academic, 1970, p. 151-164.
- PREWITT, J.M.S. Object enhancement and extraction. In: Lipkin, B.S.; Rosenfeld, A., ed. *Picture processing and psychopictorics*, New York, Academic, 1970, p.75-150.
- ROBERTS, L.G. Machine perception of three-dimensional solids. In: Tippett, J.T., et alii, ed. *Optical and electro-optical information processing*. Cambridge, MIT Press, 1965.
- SCHER, A.; VELASCO, F.R.D.; ROSENFELD, A. Some new image smoothing algorithms. *IEEE Transactions Systems, Man and Cybernetics*, SMS-10-10(3):153-158, 1980.
- TOMITA, F.; TSUJI, S. Extraction of multiple regions by smoothing in selective neighborhoods. *IEEE Transactions Systems, Man and Cybernetics*, SMC-7(2):107-109, 1977.



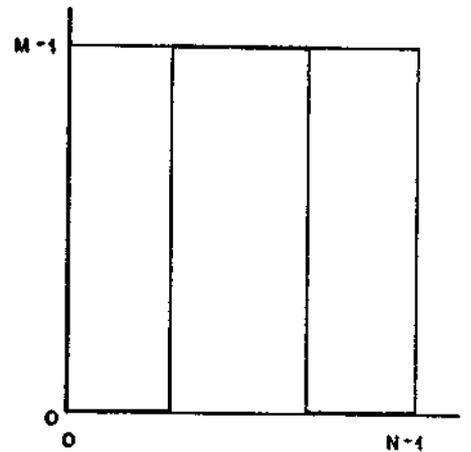
(a)



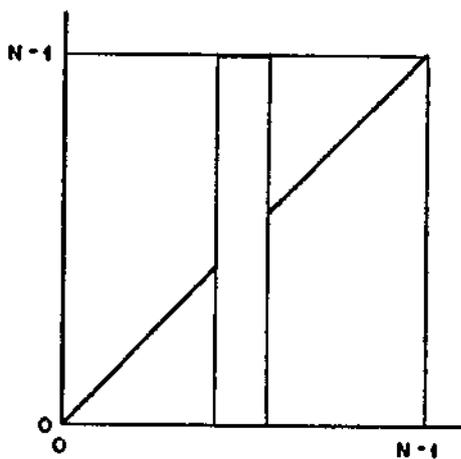
(b)



(c)



(d)



(e)

Fig. 3.1 - Mapeamento direto de n̄veis de cinza.

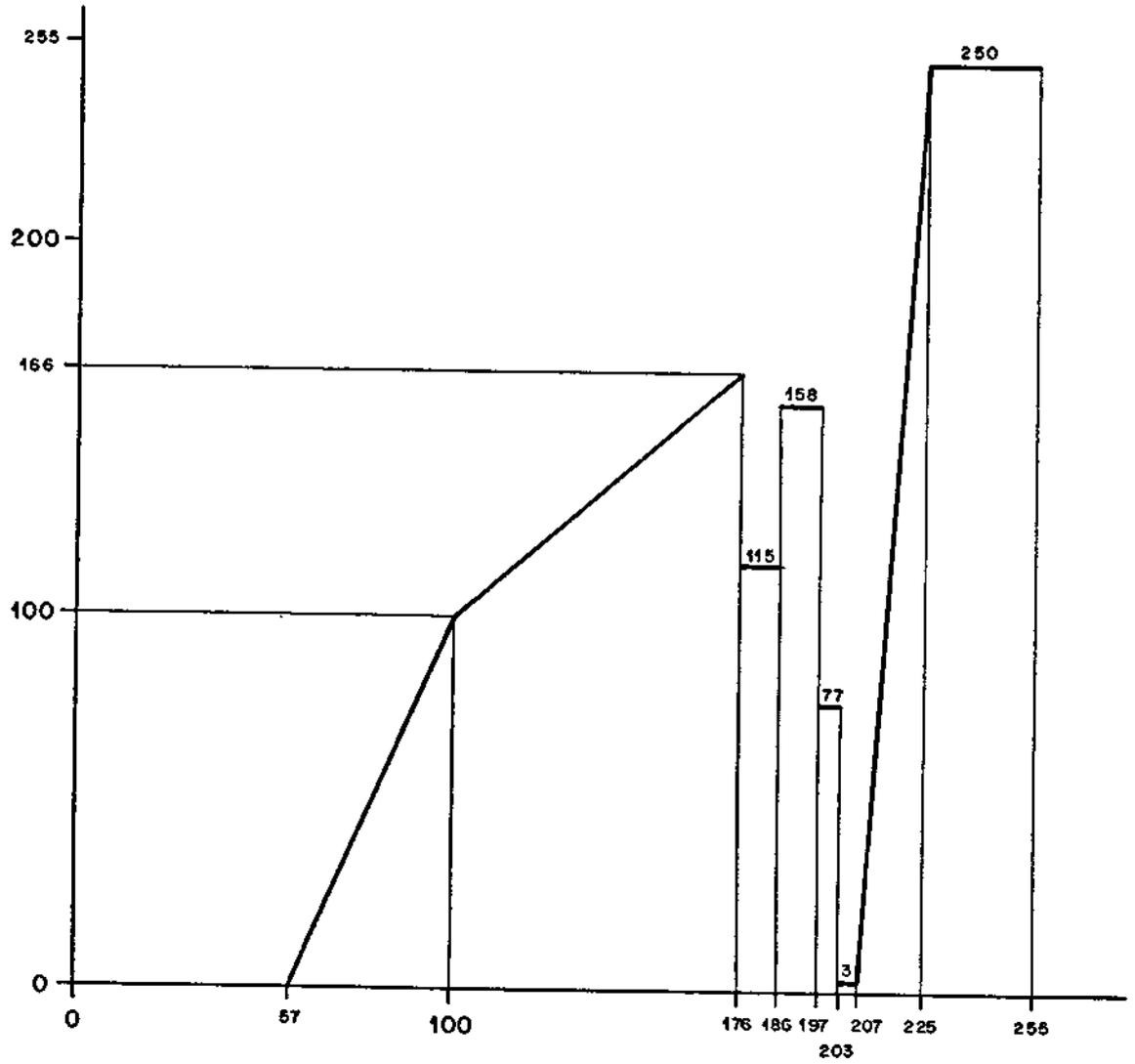
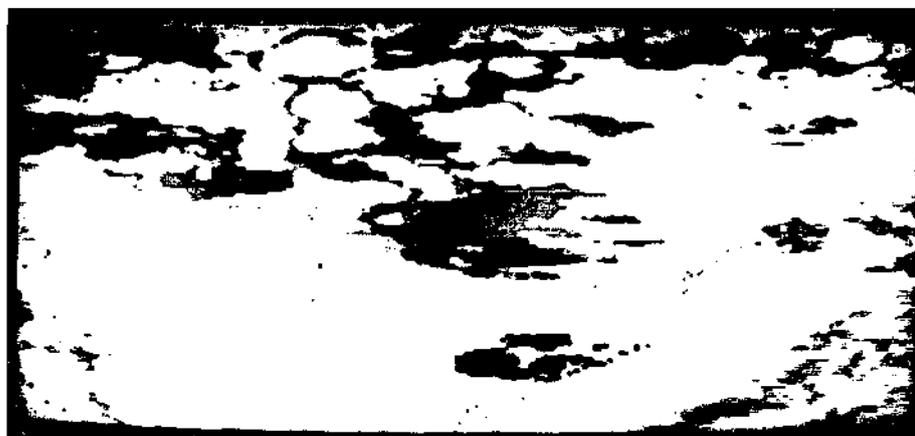


Fig. 3.2 - Curva MB.



(a)



(b)

Fig. 3.3 - Imagem da banda infravermelha do satélite SMS(a); imagem realçada pela curva MB(b).

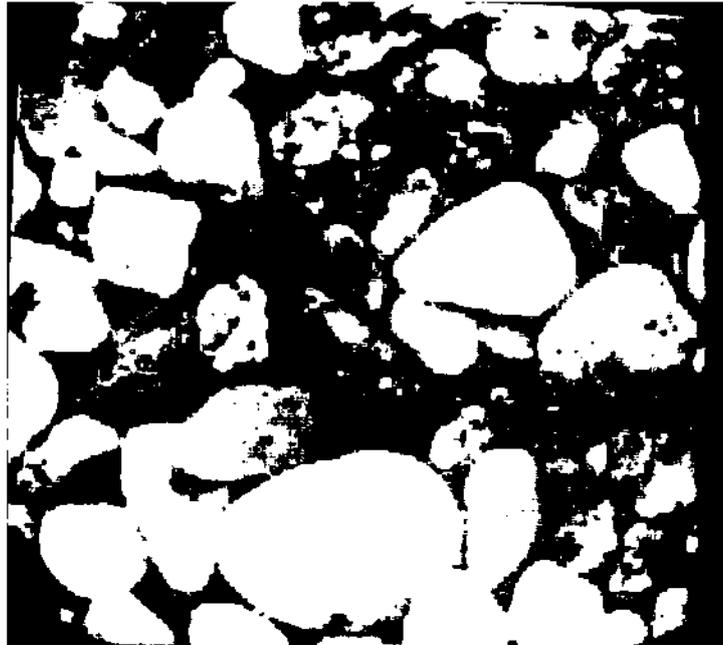


(a)

0-	43	*
0-	0	
16-	0	
24-	0	
32-	0	
40-	0	
48-	0	
56-	0	
64-	0	
72-	0	
80-	0	
88-	1	*
96-	200	***
104-	2886	*****
112-	3552	*****
120-	2563	*****
128-	2228	*****
144-	954	*****
152-	496	*****
160-	440	*****
168-	345	****
176-	534	*****
184-	398	***
192-	115	**
200-	2	*
208	0	
216-	0	
224-	0	
232-	0	
240-	0	
248-	37	*

(b)

Fig. 3.4 - Imagem (a) e histograma (b).



(a)

0-	0	
10-	98	**
20-	156	***
30-	905	*****
40-	2455	*****
50-	3362	*****
60-	2741	*****
70-	1658	*****
80-	1224	*****
90-	1261	*****
100-	1025	*****
110-	756	*****
120-	476	*****
130-	195	***
140-	42	*
150-	9	*
160-	7	*
170-	6	*
180-	2	*
190-	0	
200-	1	*
210-	3	*
220-	0	
230-	0	
240-	1	*
250-	1	*

(b)

Fig. 3.5 - Imagem (a) e histograma (b).

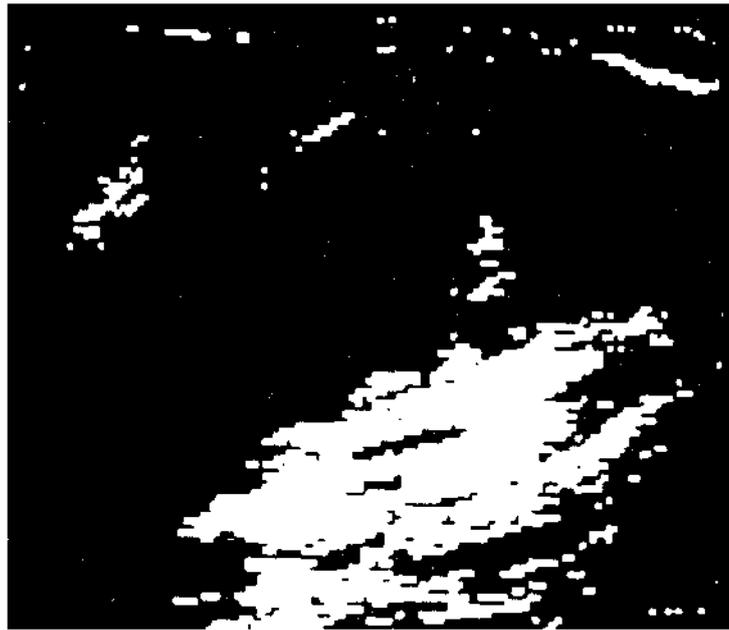
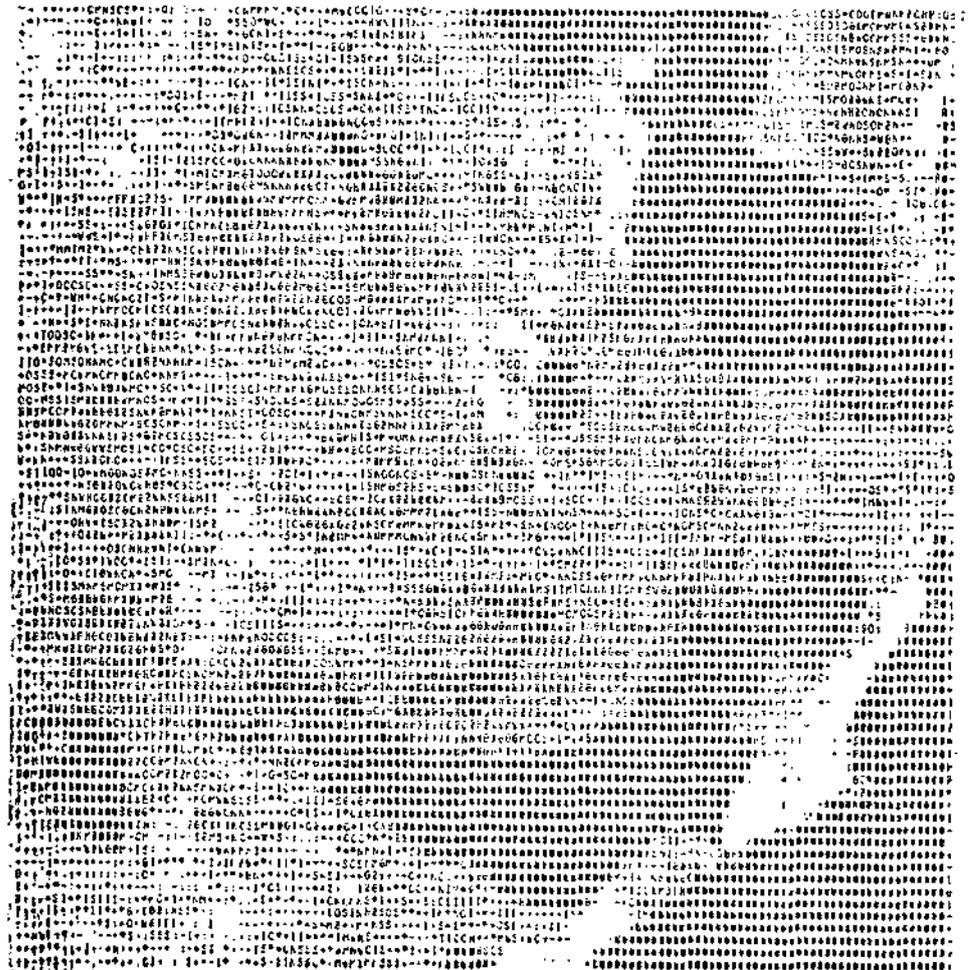


Fig. 3.6 - Imagem limiarizada.



(b)

Fig. 3.7 - Imagem original (a) e equalizada (b).

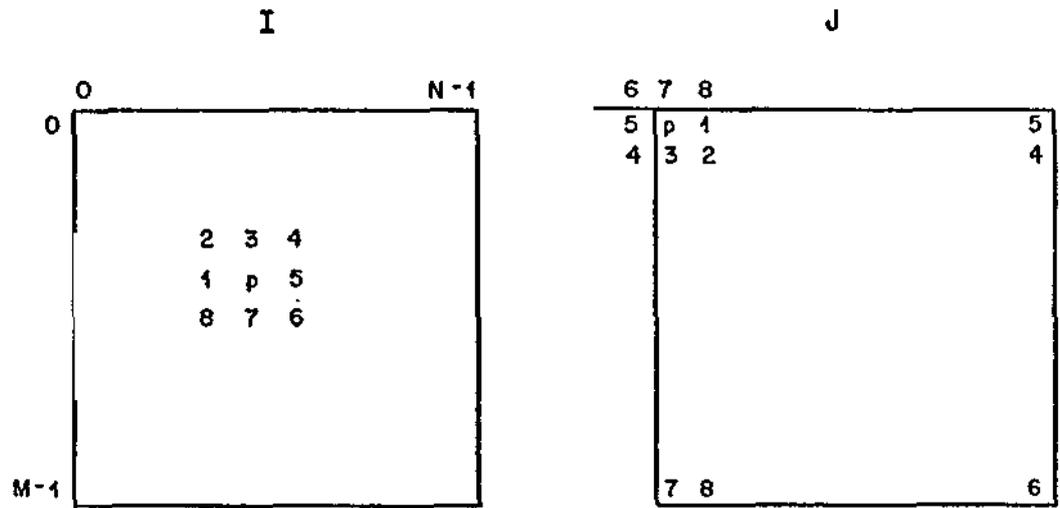


Fig. 3.8 - Convolução de duas imagens digitais I e J.

	0	1	2	3	4	5
0	a	b	c	d	c*	b*
1	e	g	h	i	j	l
2	f	m	n	o	p	q
3	f*	q*	p*	o*	n*	m*
4	e*	l*	j*	i*	h*	g*

Fig. 3.9 - Valores conjugados na transformada de Fourier de uma imagem real 5 x 6.

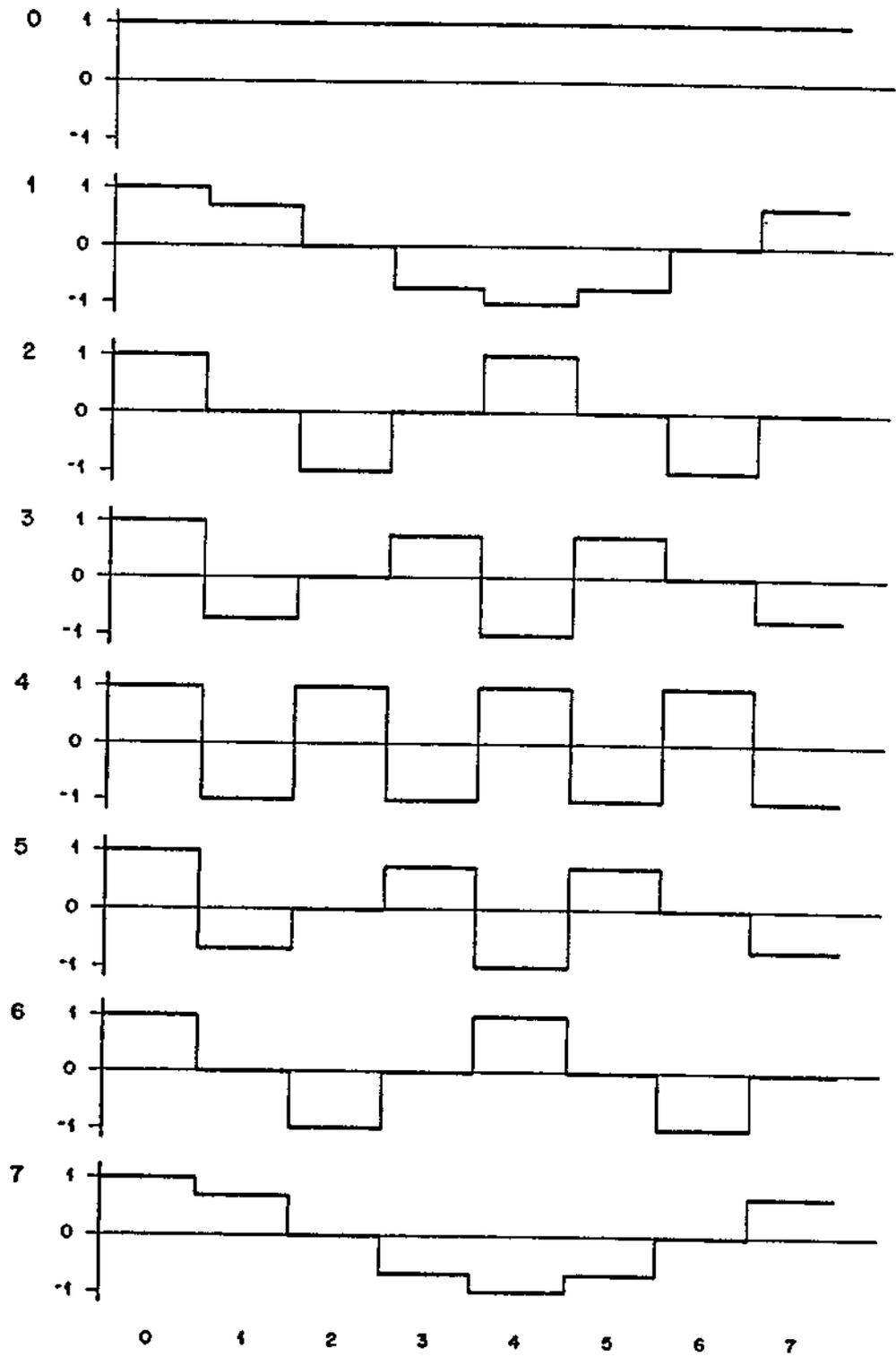


Fig. 3.10 - Componente cossenoidal da função B.

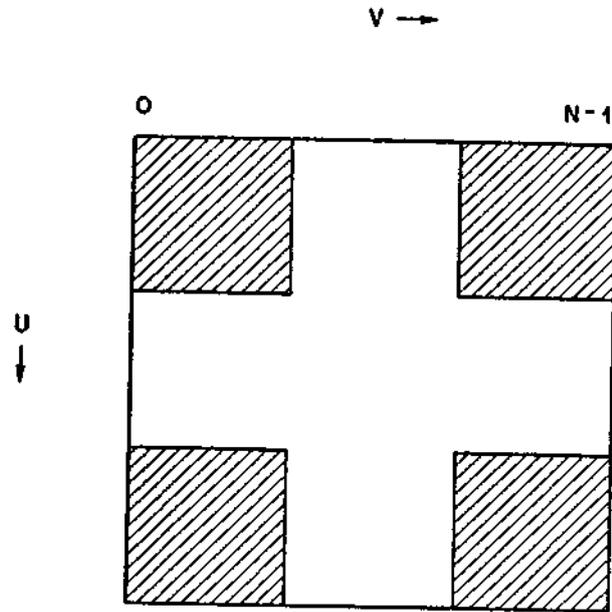


Fig. 3.11 - Regiões de baixas frequências da Transformada de Fourier.

IMAGEM DE ENTRADA								
	0	1	2	3	4	5	6	7
0	0.110	0.110	0.000	0.000	0.000	0.000	0.000	0.110
1	0.110	0.110	0.000	0.000	0.000	0.000	0.000	0.110
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	0.110	0.110	0.000	0.000	0.000	0.000	0.000	0.110
TRANSFORMADA DE FOURIER								
PARTE REAL								
	0	1	2	3	4	5	6	7
0	0.990	0.797	0.330	-0.137	-0.330	-0.137	0.330	0.797
1	0.797	0.641	0.266	-0.110	-0.266	-0.110	0.266	0.641
2	0.330	0.266	0.110	-0.046	-0.110	-0.046	0.110	0.266
3	-0.137	-0.110	-0.046	0.019	0.046	0.019	-0.046	-0.110
4	-0.330	-0.266	-0.110	0.046	0.110	0.046	-0.110	-0.266
5	-0.137	-0.110	-0.046	0.019	0.046	0.019	-0.046	-0.110
6	0.330	0.266	0.110	-0.046	-0.110	-0.046	0.110	0.266
7	0.797	0.641	0.266	-0.110	-0.266	-0.110	0.266	0.641

Fig. 3.13 - Transformada de Fourier do filtro do tipo média.

IMAGEM DE ENTRADA								
	0	1	2	3	4	5	6	7
0	5.000	-1.000	0.000	0.000	0.000	0.000	0.000	-1.000
1	-1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
7	-1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

TRANSFORMADA DE FOURIER								
PARTE REAL								
	0	1	2	3	4	5	6	7
0	1.000	1.586	3.000	4.414	5.000	4.414	3.000	1.586
1	1.586	2.172	3.586	5.000	5.586	5.000	3.586	2.172
2	3.000	3.586	5.000	6.414	7.000	6.414	5.000	3.586
3	4.414	5.000	6.414	7.828	8.414	7.828	6.414	5.000
4	5.000	5.586	7.000	8.414	9.000	8.414	7.000	5.586
5	4.414	5.000	6.414	7.828	8.414	7.828	6.414	5.000
6	3.000	3.586	5.000	6.414	7.000	6.414	5.000	3.586
7	1.586	2.172	3.586	5.000	5.586	5.000	3.586	2.172

Fig. 3.14 - Transformada de Fourier do filtro diferença.

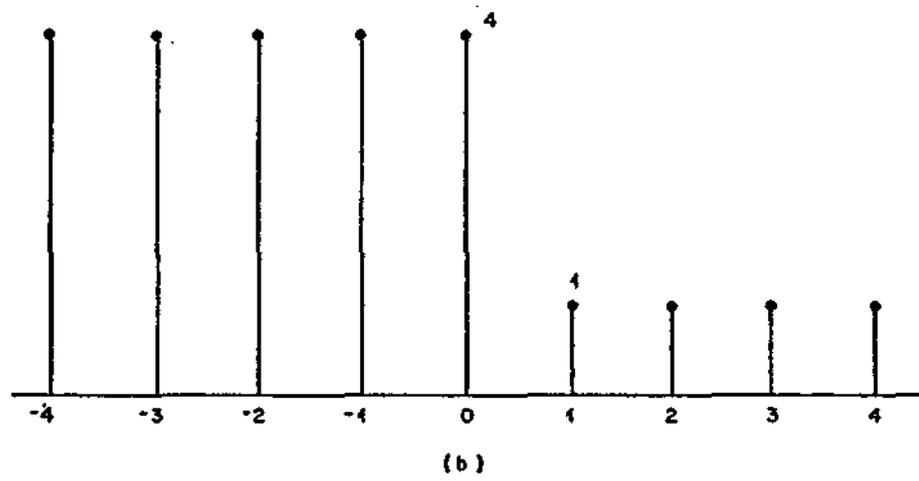
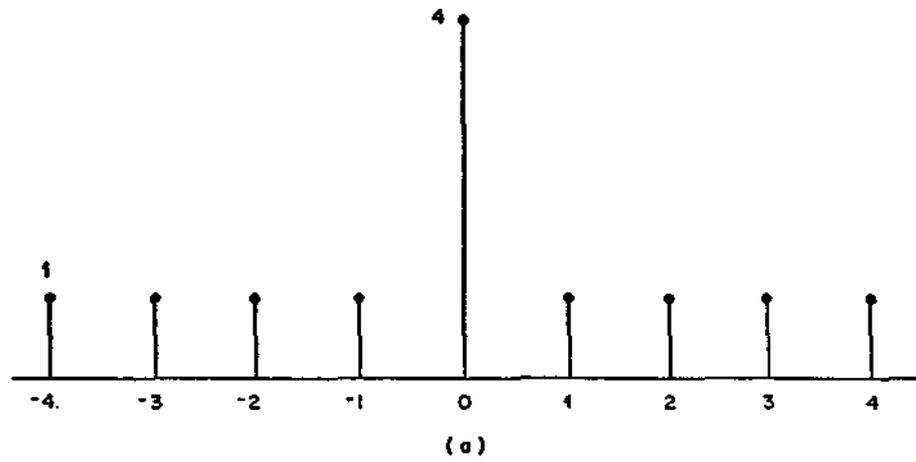


Fig. 3.18 - Sinais ruído (a) e degrau (b).

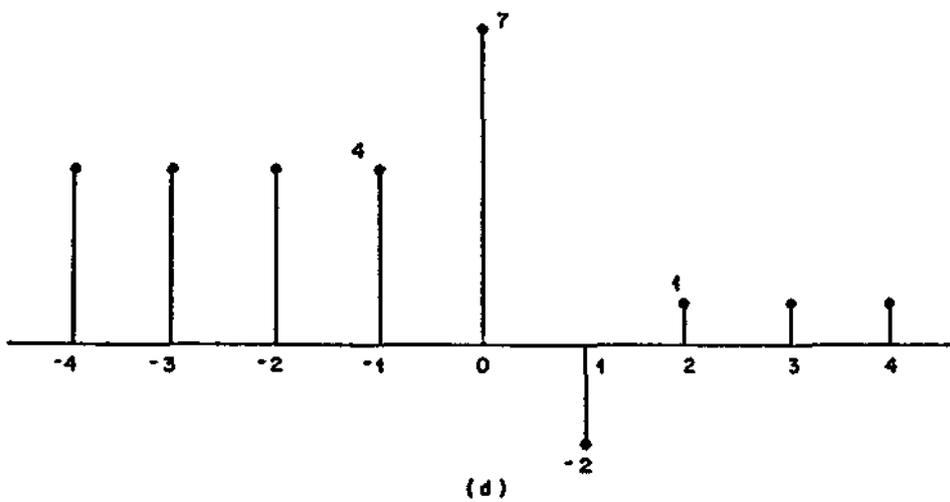
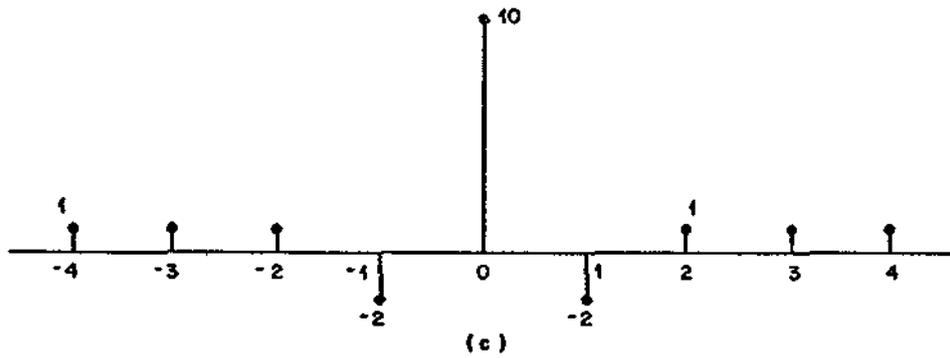
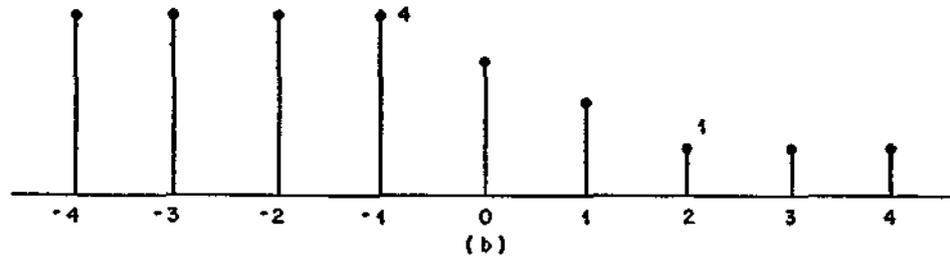
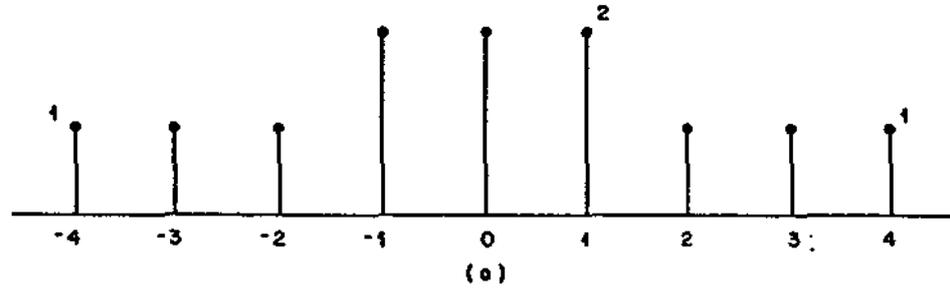
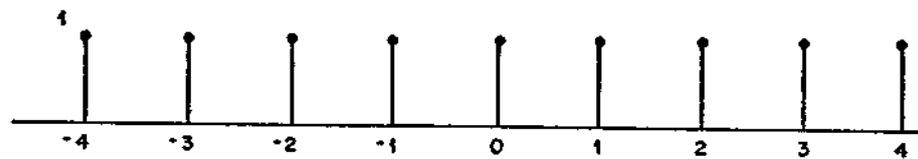
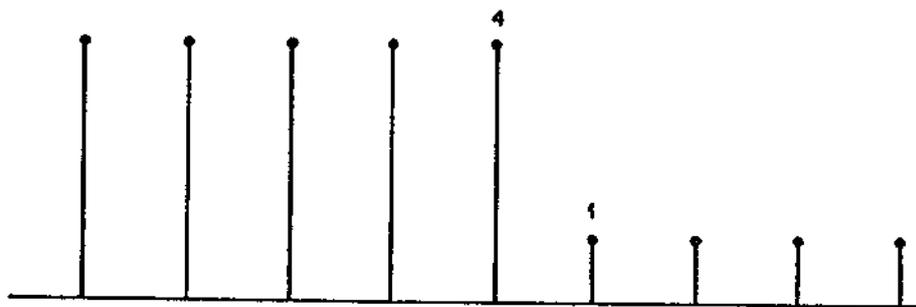


Fig. 3.19 - Efeitos de filtros lineares ao ruído e à borda: (a) efeitos de atenuação; (b) borrimento da borda; (c) ampliação do ruído; (d) realce da borda.



(a)



(b)

Fig. 3.20 - Efeito do filtro da mediana sobre o ruído e sobre o degrau.

CAPÍTULO 4

RESTAURAÇÃO

De um modo geral as técnicas de realce de imagens têm um caráter ad hoc porque frequentemente baseiam-se num critério subjetivo de avaliação, o qual é difícil de ser modelado matematicamente no estágio atual de conhecimento sobre o sistema visual humano. Por outro lado, os processos de restauração e reconstrução de imagens podem ser mais formalizadas, pois se baseiam em critérios de avaliação objetivos. Todavia, mesmo que esses problemas admitam uma solução formal, a implementação dos algoritmos de solução no computador pode ser bastante dificultada pelo esforço computacional de lidar com problemas bidimensionais.

4.1 - MODELOS DE DEGRADAÇÃO

O problema de restauração consiste na tarefa de estimar uma imagem que sofreu um processo de degradação, que envolve algum tipo de espalhamento da luz e contaminação por ruído. Estes problemas ocorrem em imagens obtidas por sensores próximos de seu limite de resolução como em Astronomia, Microscopia Eletrônica, imageamento por satélites militares de alta resolução, Radiografia Médica, ou ainda em condições adversas (movimento ou turbulência). De modo geral, o problema de restauração consiste em recuperar a imagem original da melhor maneira possível.

Para que o problema se torne matematicamente tratável, é preciso efetuar algumas simplificações. 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. Esta última simplificação pode ser bastante restritiva quando se verifica que, por exemplo, uma das causas frequentes de presença de ruído em imagens advém de granulação fotográfica que normalmente é dependente do sinal e com efeito multiplicativo. Todavia, o modelo simplificado, além de tratável matematicamente, tem-se revelado satisfatório do ponto de vista experimental em muitas situações. Desta maneira, o espalhamento da luz 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 seguinte forma:

$$g(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha,\beta) h(x,\alpha,y,\beta) d\alpha d\beta + r(x,y). \quad (4.1)$$

A Figura 4.1 ilustra o modelo simplificado de degradação.

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

$$\begin{aligned} g(x,y) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\alpha,\beta) h(x-\alpha,y-\beta) d\alpha d\beta + r(x,y) = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x-\alpha,y-\beta) h(\alpha,\beta) d\alpha d\beta + r(x,y). \end{aligned} \quad (4.2)$$

Na ausência de ruído e tomando a Transformada de Fourier de ambos os lados da Equação 4.2, verifica-se que:

$$G(u,v) = F(u,v) \cdot H(u,v), \quad (4.3)$$

onde $H(u,v)$ é a chamada função de transferência do sistema linear que caracteriza o espalhamento da luz.

Em algumas situações é possível obter um modelo físico para o processo de degradação. Considere-se, por exemplo, a situação em que uma imagem é obtida durante um certo tempo de exposição (como numa câmara fotográfica) e, durante esse tempo, ocorre um movimento relativo entre a câmara e a cena. Assim sendo, a imagem obtida $g(x,y)$ será dada por:

$$g(x,y) = \int_{-T/2}^{+T/2} f(x-x_0(t), y-y_0(t)) dt, \quad (4.4)$$

onde T é a duração da exposição e $x_0(t)$ e $y_0(t)$ são as componentes nas direções x e y do deslocamento da cena, respectivamente. Obtendo a Transformada de Fourier de (4.4), tem-se:

$$G(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[\int_{-T/2}^{+T/2} f(x-x_0(t), y-y_0(t)) dt \right] \cdot e^{-j2\pi(ux + vy)} dx dy. \quad (4.5)$$

Invertendo-se a ordem de integração na Equação 4.5, resulta em:

$$G(u,v) = \int_{-T/2}^{+T/2} \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x-x_0(t), y-y_0(t)) e^{-j2\pi(ux + vy)} dx dy \right] dt. \quad (4.6)$$

O termo dentro dos colchetes é reconhecido como a Transformada de Fourier da função deslocada $f(x-x_0(t), y-y_0(t))$, e usando a propriedade de deslocamento daquela transformada, obtêm-se:

$$\begin{aligned} G(u,v) &= \int_{-T/2}^{+T/2} F(u,v) e^{-j2\pi(ux_0(t) + vy_0(t))} dt \\ &= F(u,v) \int_{-T/2}^{+T/2} e^{-j2\pi(ux_0(t) + vy_0(t))} dt. \end{aligned} \quad (4.7)$$

Comparando-se as Equações 4.3 e 4.7, concluiu-se que:

$$H(u,v) = \int_{-T/2}^{+T/2} e^{-j2\pi(ux_0(t) + vy_0(t))} dt. \quad (4.8)$$

No caso simplificado de um movimento retilíneo uniforme na direção x , $x_0(t) = Vt$, $y_0(t) = 0$ e a integração pela Equação 4.8 fornece:

$$H(u,v) = \frac{\text{sen}(\pi UV T)}{\pi UV} . \quad (4.9)$$

Pela Transformada Inversa de Fourier da Expressão 4.9, resulta a função de resposta impulsiva bidimensional (também chamada Função de Espalhamento Pontual - F.E.P.) do sistema linear que caracteriza a degradação:

$$h(x,y) = \frac{1}{V^2 T} \text{rect}\left(\frac{x}{VT}\right), \quad (4.10)$$

$$\text{onde: } \text{rect}(x) = \begin{cases} 1 & \text{se } 0 < x < 1. \\ 0 & \text{fora} \end{cases} \quad (4.11)$$

No caso de turbulência atmosférica, Hufnagel e Stanley (1964) mostraram que a função de transferência para longas exposições pode ser aproximada por uma função do tipo $\exp(-c(u^2 + v^2)^{5/6})$. Para exposições curtas, entretanto, o problema torna-se mais difícil, pois a função de espalhamento tende a ser estocástica (Slepian, 1967).

4.2 - DETERMINAÇÃO EXPERIMENTAL DA FUNÇÃO DE ESPALHAMENTO PONTUAL

Se o fenômeno físico que causa a degradação é muito complexo para ser modelado matematicamente, deve-se recorrer a determinações experimentais da Função de Espalhamento Pontual (F.E.P.), a partir da própria imagem degradada.

4.2.1 - Fonte Pontual

Em alguns casos é razoável admitir que a cena contenha fontes de luz que possam ser consideradas pontuais, isto é, aproximem-se da função delta de Dirac. Tal é o caso de imagens astronômicas em que a imagem de uma estrela seria uma boa aproximação da F.E.P..

4.2.2 - Fonte Retilínea

Se a cena original contém, objetos com retas, é possível, sob certas condições, determinar a F.E.P. a partir da imagem degradada.

Suponha-se, por exemplo, que na cena exista uma reta paralela ao eixo y . Este objeto pode ser descrito pela função $f(x,y) = \delta(x)$. Utilizando a Equação 4.2 na ausência de ruído, tem-se que a Função de Espalhamento Retilínea (F.E.R.) pode ser dada por:

$$h_R(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(\alpha) h(x-\alpha, y-\beta) d\alpha d\beta. \quad (4.12)$$

Utilizando a propriedade de esquadramento da função delta, na integração em α , obtêm-se

$$h_R(x,y) = \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} \delta(\alpha) h(x-\alpha, y-\beta) d\alpha \right] d\beta, \quad (4.13)$$

$$h_R(x,y) = \int_{-\infty}^{+\infty} h(x, y-\beta) d\beta \quad (4.14)$$

ou, por uma mudança de variável de integração dada por $y' = y - \beta$,

$$h_R(x,y) = \int_{-\infty}^{+\infty} h(x, y') dy', \quad (4.15)$$

ou seja, a resposta do sistema degradante a uma reta na direção y é a integral da F.E.P. ao longo dessa mesma direção. Deve-se notar que $h_R(x,y)$ é função apenas de x e, portanto, pode-se denotar $h_R(x,y) = h_R(x)$.

Tomando agora a Transformada de Fourier unidimensional de $h_R(x)$, tem-se:

$$\begin{aligned}
 H_R(w) &= \int_{-\infty}^{+\infty} h_R(x) e^{-j2\pi ux} dx = \\
 &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} h(x,y) dy \right] e^{-j2\pi ux} dx \\
 &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x,y) e^{-j2\pi(ux + vy)} dx dy \Big|_{v=0}
 \end{aligned} \tag{4.16}$$

Obtêm-se, portanto:

$$H_R(u) = H(u,v) \Big|_{v=0}, \tag{4.17}$$

ou seja, a Transformada de Fourier da F.E.R. para uma reta na direção y é a função de transferência do sistema degradante calculada no eixo u . Pode-se mostrar, em geral, que a Transformada de Fourier de uma reta orientada num ângulo θ com o eixo x resulta no plano complexo $u-v$ de Fourier, em valores de $H(u,v)$, ao longo de uma reta com inclinação $\theta + 90^\circ$ (Rosenfeld and Kak, 1982), conforme ilustra a Figura 4.2.

Se retas em várias direções são disponíveis, pode-se a partir das correspondentes Transformadas de Fourier unidimensionais obter, por interpolação, uma aproximação razoável de $H(u,v)$ e, conseqüentemente, de $h(x,y)$ por uma Transformada Inversa de Fourier. Deve-se notar que este problema tem grande semelhança com o problema de reconstrução de uma imagem por suas projeções unidimensionais em várias direções, o que será discutido no próximo capítulo.

Se há razões para admitir que a F.E.P. é circularmente simétrica, $H(u,v)$ também o será e basta o conhecimento da Transformada de Fourier em uma única direção para determinar $h(x,y)$.

4.2.3 - Fonte de Borda Retilínea

Se uma cena contém bordas retilíneas em várias direções, é possível também estimar a F.E.P. Considere-se, por exemplo, uma borda ao longo do eixo y , que pode ser representada por uma função degrau $U(x)$, tal que $U(x) = 1$ para $x \geq 0$ e $U(x) = 0$ para $x < 0$. Seja $h_B(x,y)$ a imagem desse degrau. Pela Equação 4.2, na ausência de ruído, tem-se

$$h_B(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\alpha,\beta) U(x-\beta) d\alpha d\beta = h_B(x), \quad (4.18)$$

uma vez que o resultado é independente de y . Tomando a derivada de $h_B(x)$, trocando a ordem de diferenciação e integração e observando que a derivada da função degrau é a função delta de Dirac, tem-se que:

$$\begin{aligned} \frac{dh_B(x)}{dx} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\alpha,\beta) \frac{d}{dx} [U(x-\beta)] d\alpha d\beta = \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(\alpha,\beta) \delta(x-\beta) d\alpha d\beta = \\ &= \int_{-\infty}^{+\infty} h(x,\beta) d\beta. \end{aligned} \quad (4.19)$$

Da equação 4.15 pode-se concluir que:

$$\frac{dh_B(x)}{dx} = h_R(x), \quad (4.20)$$

ou seja, a derivada da imagem de uma borda retilínea é a imagem de uma reta na mesma direção. No domínio de Fourier o resultado equivalente é da do por:

$$j2\pi u H_B(u) = H_R(u). \quad (4.21)$$

Assim sendo, a técnica de reconstrução da F.E.P. do item anterior pode ser utilizada, notando-se, porém, que o operador diferenciação tende a acentuar o ruído, o que pode mascarar os resultados.

4.3 - FILTRAGEM INVERSA

Nas condições de linearidade, invariância no espaço e ausência de ruído, a Equação 4.3 é válida; portanto, 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, entretanto, problemas com a utilização desse método, ou seja:

1) $H(u,v)$ pode ter zeros, como foi observado no caso de movimento uniforme unidirecional. Na ausência de ruído há uma indeterminação, sendo impossível recuperar $F(u,v)$ nessas frequências (observar, contudo que se $H(u,v)$ tem no máximo um número contável de zeros, é possível em princípio, recuperar $f(x,y)$ porque os valores indeterminados de $G(u,v)/H(u,v)$ não contribuirão para o processo de integração da Transformada Inversa de Fourier (Rosenfeld and Kak, 1982)).

2) Na presença de ruído, tem-se:

$$G(u,v) = H(u,v) F(u,v) + R(u,v) \quad (4.22)$$

e, pela filtragem inversa, resulta em:

$$\frac{G(u,v)}{H(u,v)} = F(u,v) + \frac{R(u,v)}{H(u,v)}. \quad (4.23)$$

Nas vizinhanças dos zeros de $H(u,v)$, o segundo termo devido ao ruído tende a predominar sobre o sinal $F(u,v)$. Uma possível solução para este problema consiste em efetuar a restauração da imagem por um filtro que pondere pouco as partes do espectro da imagem degradada nas vizinhanças dos zeros de $H(u,v)$, onde a relação sinal-ruído tende a ser baixa.

É frequente o caso de $H(u,v)$ cair rapidamente com u e v , representando portanto a degradação uma filtragem passa-baixas. Por outro lado, o termo $R(u,v)$ tende a cair muito menos rapidamente com u e v . Assim sendo, deve-se truncar o filtro restaurador nas vizinhanças da origem, ou seja, o filtro inverso pode ser dado por:

$$\begin{aligned} I(u,v) &= \frac{1}{H(u,v)} \quad p/\sqrt{u^2 + v^2} \leq W_0 \\ &= 1 \quad p/\sqrt{u^2 + v^2} > W_0, \end{aligned} \quad (4.24)$$

onde W_0 deve ser escolhido de maneira a excluir os zeros de $H(u,v)$. A Figura 4.3 (a,b,c,d) ilustra o resultado do truncamento adequado do filtro inverso nas altas frequências.

4.4 - O PROBLEMA DO MALCONDICIONAMENTO E SUA SOLUÇÃO

A instabilidade numérica que ocorre no filtro inverso devido à presença de zeros em $H(u,v)$ e do ruído $r(x,y)$ está intimamente ligada ao fato de a chamada equação integral, que caracteriza a degradação 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.

O problema do malcondicionamento torna-se mais claro quando um modelo discreto é obtido para o problema de restauração. Isto pode ser feito amostrando a imagem degradada, substituindo a integral que representa o processo de degradação dado pela Equação 4.1 por uma fórmula de quadratura e estimando então a imagem original $f(x,y)$ em pontos discretos. Assim sendo, por um rastreamento do tipo "TV" do conjunto bidimensional de elementos de imagens obtidas, chega-se ao modelo discreto dado por:

$$\underline{g} = \underline{Hf} + \underline{r}, \quad (4.25)$$

onde \underline{g} é um vetor ($M \times 1$) que representa a amostragem da imagem degradada; \underline{H} é uma matriz ($M \times N$), cujos elementos dependem da função de espalhamento $h(\dots, \dots)$, bem como dos pesos da fórmula numérica de quadratura empregada; \underline{f} é um vetor ($N \times 1$) que representa a amostragem da imagem original; e \underline{r} é um vetor ($M \times 1$), cujos elementos são componentes do ruído aditivo.

Obtém-se, assim, um modelo estatístico linear que, dependendo do uso ou não de conhecimento a priori (e de seu tipo) sobre a imagem original, conduz a diferentes esquemas de estimação da imagem original, representada por \underline{f} , a partir das amostras da imagem degradada, representada por \underline{g} .

4.4.1 - Estimador de mínimos quadrados

Quando não se utiliza nenhum conhecimento a priori sobre as componentes de vetor \underline{f} , caracterizando-se esse vetor como apenas um conjunto de N parâmetros não-aleatórios a serem estimados, utiliza-se normalmente o critério de mínimos quadrados para a estimação. Seu uso é motivado mais pelo fato de ser matematicamente tratável do que por critérios de fidelidade visual, que são difíceis de ser incorporados.

Desse modo, deve-se minimizar uma forma quadrática do tipo $E(\hat{\underline{f}}) = (\underline{g} - \underline{H}\hat{\underline{f}})^T \underline{V}^{-1} (\underline{g} - \underline{H}\hat{\underline{f}})$ pela escolha de $\hat{\underline{f}}$. Na expressão anterior, \underline{V} representa a matriz de covariância do vetor de ruído \underline{r} .

A minimização é feita impondo igual a zero as derivadas parciais de $\theta(\hat{\underline{f}})$ em relação a cada componente de $\hat{\underline{f}}$:

$$\frac{\partial \theta(\hat{\underline{f}})}{\partial (\hat{\underline{f}})} = -2\underline{H}^T \underline{V}^{-1} (\underline{g} - \underline{H}\hat{\underline{f}}) = \underline{0} \quad (4.26)$$

Se $M > N$, o inverso da matriz $(\underline{H}^T \underline{V}^{-1} \underline{H})$ pode existir (nesse caso, o sistema é denominado sobredeterminado) e o estimador $\hat{\underline{f}}$ é dado por:

$$\hat{\underline{f}} = (\underline{H}^T \underline{V}^{-1} \underline{H})^{-1} \underline{H}^T \underline{V}^{-1} \underline{g}. \quad (4.27)$$

Deve-se observar que se o ruído \bar{e} é branco, a matriz \underline{Y} é diagonal e a restauração \bar{e} é obtida por:

$$\hat{f} = (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{g} = \underline{H}^+ \underline{g} , \quad (4.28)$$

onde \underline{H}^+ representa a matriz pseudo-inversa da matriz \underline{H} .

Na situação em que $M < N$ tem-se um sistema subdeterminado e há infinitas soluções para o problema de mínimos quadrados, podendo-se mostrar (Andrews and Patterson, 1976) que a solução geral é dada por:

$$\hat{f} = (\underline{K}^{-1} \underline{H})^+ \underline{V}^{-1} \underline{g} + (\underline{I} - (\underline{K}^{-1} \underline{H})^+ \underline{K}^{-1} \underline{H}) \underline{v} , \quad (4.29)$$

onde \underline{v} é um vetor arbitrário e \underline{K} é uma matriz obtida pela fatorização da matriz \underline{V} , isto é:

$$\underline{V} = \underline{K} \underline{K}^T . \quad (4.30)$$

A solução de mínima norma é dada por:

$$\hat{f} = (\underline{K}^{-1} \underline{H})^+ \underline{K}^{-1} \underline{g} \quad (4.31)$$

e, novamente, se o ruído \bar{e} é branco, a solução é obtida pela matriz pseudo-inversa \underline{H}^+ . No caso em que o posto de \underline{H} seja dado pelo número de linhas (M), essa solução pode ser expressa por:

$$\hat{f} = \underline{H}^T (\underline{H} \underline{H}^T)^{-1} \underline{g} . \quad (4.32)$$

Deve-se observar que a solução obtida pelo critério de mínimos quadrados coincide com aquela formulada pelo critério de máxima verossimilhança quando o ruído de observação é gaussiano (Sage and Melsa, 1971).

O problema de malcondicionamento da restauração é expresso pelo fato de que pequenas perturbações no vetor de observações \underline{g} , causadas pelo ruído \underline{r} , causam grandes variações na estimativa de \underline{f} . O grau de malcondicionamento do problema de restauração pode ser medido pelo chamado \underline{n}

mero de condição do modelo de regressão estatística envolvido (Mascarenhas and Pratt, 1975). Por simplificação, considere-se o caso de um sistema so bredeterminado (isto é, o posto da matriz H é dado pelo número de colunas) com ruído branco. A perturbação no estimador $\hat{\underline{f}}$ devido ao ruído é dada por:

$$\underline{\Delta \hat{f}} = \underline{H}^T \underline{r} = (\underline{H}^T \underline{H})^{-1} \underline{H}^T \underline{r}. \quad (4.33)$$

Pode-se mostrar (Rust and Burrus, 1972) que o erro relativo de estimação $\frac{\|\underline{\Delta \hat{f}}\|}{\|\underline{f}\|}$ está sujeito ao limite superior, ou seja:

$$\frac{\|\underline{\Delta \hat{f}}\|}{\|\underline{f}\|} < \|\underline{H}^+\| \|\underline{H}\| \frac{\|\underline{r}\|}{\|\underline{g}\|}. \quad (4.34)$$

O produto $\|\underline{H}^+\| \|\underline{H}\|$ é denominado número de condição de H e pode ser computado pela razão entre o maior e o menor valor singular de H (raiz quadrada do valor próprio de $\underline{H}^T \underline{H}$). Claramente, quanto maior o número de condição, maior será a sensibilidade do estimador às perturbações do ruído.

No caso de um sistema sobredeterminado, a forma quadrática a ser minimizada representa um parabolóide e o alongamento do elipsóide que resulta no plano x estabelecendo a igualdade da forma quadrática a uma constante está diretamente ligado ao número de condição do modelo linear de restauração. Quanto mais esse elipsóide se afasta do formato de uma esfera, mais malcondicionado será o modelo de restauração. De fato, pode-se mostrar que o número de condição pode ser dado pela razão entre o maior e o menor eixo principal desse elipsóide.

No caso de um sistema subdeterminado, o parabolóide degenera-se para o formato de uma calha e existem infinitas soluções para a minimização da forma quadrática. Atinge-se assim, de certa forma, o caso limite de alongamento do elipsóide.

4.4.2 - Estimador de Mínimos Quadrados com Restrição Linear de Desigualdade - Conhecimento a Priori Determinístico

O problema de malcondicionamento pode ser contornado com a intervenção de conhecimento a priori, compensando assim a incerteza advinda das observações devida à quase singularidade da matriz $(\underline{H}^T \underline{V}^{-1} \underline{H})$ no caso de um sistema sobredeterminado. Esse conhecimento pode ser determinístico, estabelecendo-se por exemplo, limites inferiores e superiores nos valores dos elementos de imagens a ser estimados (Mascarenhas and Pratt, 1975). Um limite inferior natural é zero, uma vez que os elementos de imagem representam valores de iluminância. O problema de mínimos quadráticos recai então num algoritmo de programação quadrática. A carga computacional envolvida, entretanto, pode ser bastante alta quando se lida com vetores e matrizes de grande dimensionalidade, os quais resultam do problema espacial. A utilização de métodos iterativos para a solução de problemas de programação quadrática de grande dimensões, como aqueles propostos para reconstrução de imagens (a ser abordada no próximo capítulo) parece ser uma área promissora aliviando os requerimentos de memória ao custo de um maior tempo de processamento (Censor, 1981).

4.4.3 - Estimador de Mínimos Quadrados com Restrição Quadrática de Igualdade - Técnicas de Suavização e Regularização

Os métodos de suavização e regularização (Philips, 1962; Twomey, 1963, Tikonov, 1963) também têm sido utilizados para superar o problema de malcondicionamento associado à restauração de imagens. Basicamente, estes métodos minimizam uma expressão quadrática do tipo $\underline{\tilde{f}}^T \underline{Q} \underline{\tilde{f}}$ que mede a suavidade da solução sujeita a uma restrição de igualdade na norma de vetor residual, dada por: $\| \underline{g} - \underline{H} \underline{\tilde{f}} \|^2 = e$.

Quando se impõe esta restrição, procura-se um ponto de mínimo da expressão lagrangiana:

$$W(\underline{\tilde{f}}, \lambda) = \underline{\tilde{f}}^T \underline{Q} \underline{\tilde{f}} + \lambda [(\underline{g} - \underline{H} \underline{\tilde{f}})^T (\underline{g} - \underline{H} \underline{\tilde{f}}) - e] \quad (4.35)$$

Tomando as derivadas com respeito a $\underline{\tilde{f}}$ e a λ e igualando-as a zero, obtêm-se:

$$\frac{\partial W}{\partial \underline{\hat{f}}} = 2\underline{Q} \underline{\hat{f}} + \lambda [-2 \underline{H}^T \underline{g} + 2\underline{H}^T \underline{H}^T \underline{\hat{f}}] = 0, \quad (4.36)$$

$$\frac{\partial W}{\partial \lambda} = (\underline{g} - \underline{H} \underline{\hat{f}})^T (\underline{g} - \underline{H} \underline{\hat{f}}) - e = 0. \quad (4.37)$$

Da Equação 4.36 segue-se que:

$$(\underline{Q} + \lambda \underline{H}^T \underline{H}) \underline{\hat{f}} = \lambda \underline{H}^T \underline{g} \quad (4.38)$$

Dividindo-se por λ ambos os membros da Equação 4.38, pode-se obter a solução $\underline{\hat{f}}$ dada por:

$$\underline{\hat{f}} = (\underline{H}^T \underline{H} + \frac{1}{\lambda} \underline{Q})^{-1} \underline{H}^T \underline{g}, \quad (4.39)$$

onde λ é escolhido iterativamente de modo que a restrição de igualdade se ja satisfeita.

Se, por outro lado, for considerado o problema de estimar \underline{f} por um critério de mínimos quadrados, isto é, minimizar $(\underline{g} - \underline{H} \underline{\hat{f}})^T (\underline{g} - \underline{H} \underline{\hat{f}})$ tal que uma restrição quadrática de igualdade $\underline{\hat{f}}^T \underline{Q} \underline{\hat{f}} = d$ na medida de suavidade da solução seja satisfeita, é fácil verificar que uma solução do mesmo tipo dado pela Equação 4.39 é obtida. Desse modo, os métodos de regularização ou suavização podem ser encarados como métodos iterativos para resolver um problema de mínimos quadrados com restrição quadrática de igualdade. Para $\lambda \rightarrow \infty$, a solução de mínimos quadrados é obtida. Para um valor finito de λ a matriz $\frac{1}{\lambda} \underline{Q}$ tenderá a evitar o problema de malcondicionamento que advém da quase-singularidade da matriz $\underline{H}^T \underline{H}$.

A escolha da matriz \underline{Q} deve recair sobre uma medida adequada de suavidade da solução. No caso unidimensional, uma possível medida seria expressa pela segunda derivada da função $f(x)$, que poderia ser aproximada por:

$$p(x,y) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (4.44)$$

É interessante observar que os métodos de regularização e sua vização estão relacionados aos chamados estimadores de mínimos quadrados obtidos por regressão de crista ("ridge regression") (Marquardt, 1970), desenvolvidos na literatura estatística. Do ponto de vista estatístico, quando o sistema é sobredeterminado, sabe-se que o estimador de mínimos quadrados é o de mínima variância sem viés. (Lewis and Odell, 1971). A adição da matriz $\frac{1}{\lambda} Q$ à matriz $(H^T H)$ provoca uma redução das variâncias das componentes de \hat{f} ao preço da introdução de um viés desconhecido. Quando $\lambda \rightarrow \infty$, $\hat{f} \rightarrow \underline{0}$, as variâncias de \hat{f} tendem a zero e o viés (desconhecido) é o vetor que representa a imagem original \underline{f} .

4.4.4 - Filtragem de Wiener - Conhecimento a Priori Estatístico

Outra possibilidade de restauração consistiria no uso do conhecimento a priori do tipo estatístico sobre a imagem original, recaindo, portanto, num esquema bayesiano de estimação.

O critério utilizado normalmente neste caso é o do erro médio quadrático, a despeito de se saber em analogia ao que foi mencionado na Seção 4.4.1, que tal critério não incorpora as propriedades do sistema visual humano.

Se nenhuma restrição é colocada nas operações a serem efetuadas na imagem degradada $g(x,y)$, para obter um estimador $\hat{f}(x,y)$ da imagem original $f(x,y)$, é possível demonstrar (Papoulis, 1965) que $\hat{f}(x,y)$ é o valor esperado condicional de $f(x,y)$ dado $g(x,y)$, o que, em geral, é uma função não-linear de $g(x,y)$. Além disso, esta operação envolve o conhecimento das densidades de probabilidade conjuntas das variáveis aleatórias que descrevem os campos aleatórios $g(x,y)$ e $f(x,y)$, o que é difícil de obter. Por essas razões, esse método é difícil de ser aplicado, tanto do ponto de vista teórico como de implementação computacional.

Se for feita a restrição de utilizar apenas operações lineares, ou se se dispuser de conhecimento de momentos até 2ª ordem sobre $g(x,y)$ e $f(x,y)$, ou ainda, se os campos aleatórios $g(x,y)$, $f(x,y)$ e $r(x,y)$ forem conjuntamente gaussianos, o procedimento ótimo é o chamado filtro de Wiener (Papoulis, 1975).

4.4.4.1 - Caso Contínuo

Denotando por \underline{r} o veto de coordenadas (x,y) no plano, pode-se expressar a operação linear do filtro restaurador ótimo por:

$$\hat{f}(\underline{r}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\underline{r}, \underline{r}') g(\underline{r}') d\underline{r}'. \quad (4.45)$$

Se os campos aleatórios envolvidos são homogêneos, w será função apenas da diferença $(\underline{r} - \underline{r}')$, obtendo-se:

$$\hat{f}(\underline{r}) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\underline{r} - \underline{r}') g(\underline{r}') d\underline{r}'. \quad (4.46)$$

A expressão a ser minimizada pelo critério de erro médio quadrático é, portanto, dada por:

$$\epsilon^2 = E \left\{ \left[f(\underline{r}) - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\underline{r} - \underline{r}') g(\underline{r}') d\underline{r}' \right]^2 \right\}. \quad (4.47)$$

É possível mostrar (Rosenfeld and Kak, 1982) que a função $w(\underline{r} - \underline{r}')$ que satisfaz a seguinte equação:

$$E \left\{ \left[f(\underline{r}) - \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\underline{r} - \underline{r}') g(\underline{r}') d\underline{r}' \right] g(\underline{s}) \right\} = 0 \quad (4.48)$$

para todos os vetores de posição \underline{r} e \underline{s} no plano, minimiza a Equação 4.47. Este resultado é conhecido como o princípio de ortogonalidade, que estabelece basicamente que o erro de estimação ótima é ortogonal aos dados de

observação.

Efetuada a operação de valor esperado dentro do sinal de integração, levando em conta as definições de função de autocorrelação e correlação cruzada, e após as mudanças de variáveis $\underline{r}' - \underline{s}' = \underline{\tau}'$ e $\underline{r} - \underline{s} = \underline{\tau}$, obtêm-se:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} w(\underline{\tau} - \underline{\tau}') R_{gg}(\underline{\tau}') d\underline{\tau}' = R_{fg}(\underline{\tau}). \quad (4.49)$$

A Equação 4.49 representa uma convolução bidimensional e tomando a Transformada de Fourier dos dois lados da Equação 4.49, tem-se a expressão da função de transferência do filtro ótimo:

$$W(u,v) = \frac{S_{fg}(u,v)}{S_{gg}(u,v)}, \quad (4.50)$$

onde $S_{fg}(u,v)$ e $S_{gg}(u,v)$ representam, respectivamente, as funções densidade de spectral de potência cruzada entre $f(x,y)$ e $g(x,y)$ e de $g(x,y)$ consigo mesmo. Sob a hipótese de que o sinal $f(t)$ e o ruído $r(t)$ são não-correlacionados e pelo menos um deles tem média zero, pode-se mostrar (Rosenfeld and Kak, 1982) que a Equação 4.40 pode ser expressa como:

$$\begin{aligned} W(u,v) &= \frac{H^*(u,v) S_{ff}(u,v)}{S_{ff}(u,v) |H(u,v)|^2 + S_{rr}(u,v)} \\ &= \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + \frac{S_{ff}(u,v)}{S_{rr}(u,v)}}. \end{aligned} \quad (4.51)$$

Deve-se observar que na ausência de ruído $S_{rr}(u,v) = 0$, e o filtro de Wiener torna-se o filtro inverso $1/H(u,v)$. Assim, a introdução de conhecimento a priori estatístico introduz termos que evitam os problemas numéricos devidos aos zeros de $H(u,v)$.

Se o ruído $r(x,y)$ puder ser considerado branco, o termo $S_{rr}(u,v)$ na Equação 4.41 pode ser aproximado por um valor constante. Além disso, na ausência do conhecimento de $S_{ff}(u,v)$ e $S_{rr}(u,v)$, pode-se aproximar a Equação 4.51 pela expressão:

$$W(u,v) \cong \frac{1}{H(u,v)} \frac{|H(u,v)|^2}{|H(u,v)|^2 + K}, \quad (4.52)$$

onde K é uma constante que deve representar o conhecimento a priori que se tem sobre a relação sinal-ruído (RSR), isto é, a relação entre a potência do sinal $f(x,y)$ e a do ruído $r(x,y)$.

4.4.4.2 - Caso Discreto

Partindo-se do modelo linear básico dado pela Equação 4.25, o problema da filtragem discreta de Wiener consiste na obtenção de um estimador linear \hat{f} de f , do tipo $\hat{f} = \underline{Wg}$, que minimiza a expressão:

$$Q = E \{(\underline{f} - \hat{\underline{f}})^T (\underline{f} - \hat{\underline{f}})\} = \text{traço} \{E [\underline{f} - \hat{\underline{f}}] [\underline{f} - \hat{\underline{f}}]^T\}. \quad (4.53)$$

Supondo-se que as estatísticas abaixo sejam conhecidas, têm-se:

$$E(\underline{f}) = E(\underline{r}) = \underline{0}, \quad (4.54)$$

$$E(\underline{f}\underline{f}^T) = \underline{C}, \quad (4.55)$$

$$E(\underline{r}\underline{r}^T) = \underline{V}. \quad (4.56)$$

Supondo ainda que os vetores \underline{f} e \underline{r} sejam não-correlacionados, têm-se:

$$E(\underline{g}) = \underline{0}, \quad (4.57)$$

$$E(\underline{g}\underline{g}^T) = \underline{H} \underline{C} \underline{H}^T + \underline{V}, \quad (4.58)$$

$$E(\underline{f}\underline{g}^T) = \underline{C} \underline{H}^T. \quad (4.59)$$

Substituindo uma solução do tipo $\hat{\underline{f}} = \underline{W} \underline{g}$ na Equação 4.53 e usando o princípio de ortogonalidade, obtêm-se:

$$E\{(\underline{f} - \underline{W}\underline{g})\underline{g}^T\} = \underline{0} \quad (4.60)$$

e, portanto,

$$\underline{W} \cdot E(\underline{g}\underline{g}^T) = E(\underline{f}\underline{g}^T). \quad (4.61)$$

Empregando as Equações 4.57 e 4.50 na Expressão 4.61, tem-se:

$$\hat{\underline{f}} = \underline{C}\underline{H}^T(\underline{H} \underline{C} \underline{H}^T + \underline{V})^{-1}\underline{g}. \quad (4.62)$$

A expressão acima pode ser colocada numa forma equivalente (Tewarson, 1972), ou seja:

$$\hat{\underline{f}} = (\underline{H}^T \underline{H} + \underline{C}^{-1}\underline{V})^{-1} \underline{H}^T \underline{g}. \quad (4.63)$$

Utilizando novamente o princípio de ortogonalidade, o erro médio quadrático pode ser dado por:

$$\begin{aligned} Q &= \text{traço} \{ E [(\underline{f} - \hat{\underline{f}})\underline{f}^T] \} = \\ &= \text{traço} \{ E [\underline{f}\underline{f}^T] - \underline{W} \cdot E [\underline{g}\underline{f}^T] \} = \\ &= \text{traço} \{ \underline{C} - \underline{W} \underline{H} \underline{C} \}. \end{aligned} \quad (4.64)$$

A redução de C do valor a priori traço $\{C\}$ pelo valor traço $\{WH C\}$ representa o efeito do filtro ótimo.

4.5 - PROBLEMA BIDIMENSIONAL - SOLUÇÃO NO DOMÍNIO DE FOURIER

Resolvido o problema do malcondicionamento resta, contudo, o problema da dimensionalidade das matrizes e vetores envolvidos no caso de deconvolução de imagens. Este problema foi resolvido pela utilização de técnicas de Transformada de Fourier (Hunt, 1973), quando o processo de de gradação é invariante no espaço. A seguir, faz-se uma descrição resumida do procedimento desenvolvido. Para maiores detalhes sobre a técnica, que envolve considerável manipulação de vetores e matrizes, o leitor deve consultar a referência original (Hunt, 1973) ou (Rosenfeld and Kak, 1982) ou ainda (Andrews and Hunt, 1977).

No caso unidimensional, supondo que haja N amostras no sinal a ser restaurado e J amostras na resposta impulsiva do sistema degradante, pelo processo de convolução obtêm-se $N + J - 1$ amostras no sinal degradado. Os vetores \underline{g} , \underline{f} e \underline{r} são estendidos com zeros até uma dimensão $P \geq N+J-1$, obtendo-se assim vetores \underline{g}_e , \underline{f}_e , \underline{r}_e . Forma-se também uma matriz \underline{H}_e de dimensão $P \times P$, obtida de tal modo que suas linhas são as amostras da resposta impulsiva do sistema degradante estendida com zeros até a dimensão P . A matriz \underline{H}_e tem a característica de ser circulante, isto é, suas linhas (ou colunas) são versões deslocadas (com periodicidade igual a P) umas das outras. Isto permite que \underline{H}_e seja fatorizada na forma $\underline{H}_e = \underline{W} \underline{D} \underline{W}^{-1}$, onde o elemento (k, i) de \underline{W} é do tipo $\exp[j \frac{2\pi}{M} ki]$ e \underline{D} é uma matriz diagonal cujos elementos da diagonal principal formam a Transformada Discreta de Fourier de uma linha de \underline{H}_e (multiplicadas por D). Com essa fatorização obtêm-se um modelo do tipo:

$$\underline{g}_e = \underline{W} \underline{D} \underline{W}^{-1} \underline{f}_e + \underline{r}_e, \quad (4.65)$$

ou ainda:

$$\underline{W}^{-1} \underline{g}_e = \underline{D} \underline{W}^{-1} \underline{f}_e + \underline{W}^{-1} \underline{r}_e. \quad (4.66)$$

A operação de multiplicação por \underline{W}^{-1} é equivalente a efetuar a Transformada Discreta de Fourier, o que resulta em:

$$G(u) = P H(u) F(u) + R(u) \quad (4.67)$$

$$u = 0, 1, \dots, P-1.$$

No caso bidimensional, as sequências são igualmente estendidas nas direções x (até P) e y (até Q) obtendo-se vetores \underline{g}_e , \underline{f}_e e \underline{r}_e de dimensão PQ e uma matriz \underline{H}_e de dimensão $P.Q \times P.Q$ que tem a característica de ser circulante por blocos, isto é, é constituída por submatrizes circulantes e a disposição das submatrizes na matriz \underline{H}_e também é circulante. Essa matriz pode ser diagonalizada pela Transformada Discreta de Fourier Bidimensional, o que resulta, por analogia com a Equação 4.67, em:

$$G(u,v) = PQ H(u,v) F(u,v) + R(u,v) \quad (4.68)$$

$$u = 0, \dots, P-1$$

$$v = 0, \dots, Q-1.$$

A partir da Equação 4.67 ou 4.68, obtêm-se relações escalares entre as componentes das Transformadas Discretas de Fourier.

O filtro restaurador inverso pode ser obtido diretamente da Equação 4.48 por:

$$\bar{F}(u,v) = \frac{1}{PQ} G(u,v) \quad (4.69)$$

e a imagem restaurada é obtida pela Transformada Discreta de Fourier Inversa.

Os métodos de suavização envolvem uso de $\underline{Q} = \underline{C}^T \underline{C}$, onde \underline{C} representa um operador do tipo convolução (e, portanto, invariante no espaço). Analogamente, essa matriz pode ser transformada numa matriz circulante (no caso unidimensional) ou circulante por blocos (no caso bidimensional), e a Equação 4.39 assume a forma:

$$\bar{f} = (\underline{W} \underline{D}^* \underline{D} \underline{W}^{-1} + \frac{1}{\lambda} \underline{W} \underline{E}^* \underline{E} \underline{W}^{-1})^{-1} \underline{W} \underline{D}^* \underline{W}^{-1} \underline{g}, \quad (4.70)$$

ou ainda:

$$\underline{W}^{-1} \underline{\hat{f}} = (\underline{D}^* \underline{D} + \frac{1}{\lambda} \underline{E}^* \underline{E})^{-1} \underline{D}^* \underline{W}^{-1} \underline{g}, \quad (4.71)$$

onde $\underline{D}^*(\underline{E}^*)$ indica a matriz cujos elementos são o complexo conjugado dos elementos de $\underline{D}(E)$. Assim, obtêm-se:

$$\hat{F}(u,v) = \left[\frac{H^*(u,v)}{|H(u,v)|^2 + \frac{1}{\lambda} |P(u,v)|^2} \right] G(u,v). \quad (4.72)$$

A Figura 4.4 ilustra as restaurações de uma imagem degradada por uma F.E.P. do tipo gaussiano, com ruído aditivo de distribuição uniforme, obtidas com o filtro inverso (Figura 4.4c) e com o método de mínimos quadrados com restrição quadrática de igualdade (Figura 4.4d), o que evidencia a substancial melhoria no último caso.

O filtro discreto de Wiener também pode ser implementado pelo uso de F.F.T.. O papel da matriz de suavização Q é agora exercido pelo produto $\underline{C}_f^{-1} \underline{V}$, conforme evidencia a comparação das Equações 4.39 e 4.63. Sob a hipótese de que os campos aleatórios do sinal e do ruído sejam homogêneos, as matrizes de covariância \underline{C}_f e \underline{V} do sinal e do ruído, respectivamente, serão matrizes de Toeplitz (Graybill, 1969), isto é, todos os elementos de cada subdiagonal paralela à diagonal principal serão iguais. Além disso, como a correlação entre os elementos de imagens tende a cair com a distância, tais matrizes terão elementos não-nulos numa faixa ao redor da diagonal principal. Mostra-se então (Andrews and Hunt, 1977) que \underline{C}_f e \underline{V} podem ser aproximadas por matrizes circulares por blocos que podem ser diagonalizadas pela Transformada Discreta de Fourier Bidimensional, ou seja:

$$\underline{C}_f = \underline{W} \underline{A} \underline{W}^{-1}, \quad (4.73)$$

$$\underline{V} = \underline{W} \underline{B} \underline{W}^{-1}. \quad (4.74)$$

Os elementos das matrizes diagonais \underline{A} e \underline{B} são, respectivamente, a densidade espectral de potência das sequências bidimensionais $f_e(x,y)$ e $r_e(x,y)$ que deram origem aos vetores \underline{f}_e e \underline{r}_e , dados por $S_f(u,v)$ e $S_r(u,v)$.

Resulta daí o filtro discreto de Wiener, dado por:

$$\underline{W}^{-1} \underline{\hat{f}} = (\underline{D}^* \underline{D} + \underline{A}^{-1} \underline{B})^{-1} \underline{D}^* \underline{W}^{-1} \underline{g}, \quad (4.75)$$

ou ainda:

$$\hat{F}(u,v) = \frac{H^*(u,v)}{|H(u,v)|^2 + \frac{S_r(u,v)}{S_f(u,v)}} G(u,v) \quad (4.76)$$

$$u_* = 0, 1 \dots, P-1$$

$$v = 0, 1 \dots, Q-1.$$

4.6 - OUTRAS TÉCNICAS DE RESTAURAÇÃO

Várias outros métodos de restauração têm sido propostos e são brevemente cobertos nesta seção.

O uso de técnicas de estimação não-linear representa uma abordagem para lidar com as não-linearidades associadas ao processo de gravação de imagens em filmes fotográficos, por exemplo (Hunt, 1977). O método envolve o critério de máxima estimação a posteriori e um processo iterativo de otimização pelo método do gradiente.

As técnicas de restauração recursiva têm sido objeto de grande número de investigações nos últimos anos, visando estender para duas dimensões os esquemas unidimensionais de filtragem de Kalman (Rosenfeld and Kak, 1982). Tais técnicas utilizam a representação de imagens por campos aleatórios do tipo markoviano no sentido amplo, podendo ser formulada

REFERÊNCIAS BIBLIOGRÁFICAS

- ANDREWS, H.C.; HUNT, B.R. *Digital image restoration*. Englewood Cliffs, N.J., Prentice Hall, 1977.
- ANDREWS, H.C.; PATTERSON, C.L. Outer Product Expansions and Their Uses in Digital Image Processing. *IEEE Transactions Computers*, C-25; (2): 140-148, Feb. 1976.
- CENSOR, Y. Row-Action Methods for Huge and Sparse Systems and their Applications. *SIAM Review*, 23: 44-464, 1981.
- FRIEDEN, B.R. Restoring with Maximum Likelihood and Maximum Entropy. *Journal Optical Society of America*, 62(4): 511-518, 1972.
- GONZALEZ, R.C.; WINTZ, P. *Digital image processing*. Reading, MA, Addison Wesley, 1977.
- GRAYBILL, F.A. *Introduction to matrices with applications in statistics*. Belmont, CA, Wadsworth, 1969.
- HARRIS, J.L. Image evaluation and restoration. *Journal Optical Society of America*, 56(5): 569-574, May, 1966.
- HUFNAGEL, R.E.; STANLEY, N.R. Modulation transfer function associated with image transmission through turbulent media. *Journal Optical Society of America*, 54(1): 52-61, Jan. 1964.
- HUNT, B.R. The application of constrained least squares estimation to image restoration by digital computer. *IEEE Transactions Computers*, C-22(9): 805-812, 1973.
- Bayesian methods in nonlinear digital image restoration. *IEEE Transactions Computers*, C-22(3): 219-229, 1977.
- JAYNES, E.T. On the rational of maximum entropy methods. *Proceedings of the IEEE*, 70(9): 939-952, Sept. 1982.
- LEWIS, T.D.; ODELL, P.L. *Estimation in linear models*. Englewood Cliffs, N.J., Prentice Hall, 1971.
- MARQUARDT, O.W. Generalized Inverses, Ridge Regression, Biased Linear Estimation and Nonlinear Estimation. *Technometrics*, 12: 591-612, 1970.

- MASCARENHAS, N.D.A.; PRATT, W.K. Digital image restoration under a regression model, *IEEE Transactions Circuits Systems*, C-22(3):252-266, Mar. 1975.
- PAPOULIS, A. *Probability, random variables and stochastic processes*. New York, Mc Graw-Hill, 1965.
- PHILIPS, D.L. A technique for the numerical solution of certain integral equations of the first kind. *Journal of the Association for Computing Machinery*, 9: 84-97, 1962.
- ROBBINS, G.M.; HUANG, T.S. Inverse Filtering for Linear Shift-Variant Imaging Systems. *Proceedings of IEEE*, 66(7): 862-872, July, 1972.
- ROSENFELD, A.; KAK, A.C. *Digital picture processing*. 2nd ed. New York, NY, Academic, 1982.
- RUST, B.W.; BURRUS, W.R. *Mathematical programming and the numerical solution of linear equations*. New York, NY, American Elsevier, 1972.
- SAGE, A.P.; MELSA, J.L. *Estimation theory with applications to communications and control*. New York, NY, Mc Graw-Hill, 1971.
- SAWCHUCK, A.A. Space-variant motion degradation and restoration. *Proceedings of the IEEE*, 60(7):854-861, July 1972.
- Space-variant image restoration by coordinate transformations. *Journal Optical Society of America*, 64(2): 138-144, Feb. 1974.
- SLEPIAN, D. Linear least squares filtering of distorted images. *Journal Optical Society America*, 57(7): 918-922, 1967.
- SLEPIAN, D.; POLLAK, H.O. Prolate spheroidal wave functions, Fourier analysis and uncertainty - I. *Bell System Technical Journal*, 40(1): 43-62, Jan. 1961.
- STULLER, J.A.; KURZ, B. Two Dimensional Markov Representations of Sampled Images. *IEEE Transactions on Communications*, 24(10): 1148-1152, Oct. 1976.
- TEWARSON, R.P. Solution of linear equations in remote sensing and picture reconstructions. *Computing*, 10: 221-230, 1972.

- TIKONOV, A.N. Regularization of incorrectly posed problems. *Soviet Math.*, 4(6): 1624-1627.
- TWOMEY, S. On the numerical solution of fredholm integral of the first kind by the inversion of the linear system produced by quadrature. *Journal of Association of Computing Machinery*, 10: 97-101, 1963.
- YOULA, D.C. Restoration by the Method of Alternating Orthogonal Projections. In: Wilde, C.O.; Barret, E., ed. *Image science mathematics*, North Hollywood, Western Periodicals, 1976, p. 176-185.

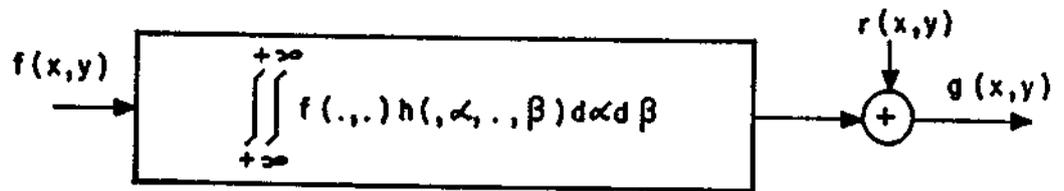


Fig. 4.1 - Modelo de degradação da imagem.

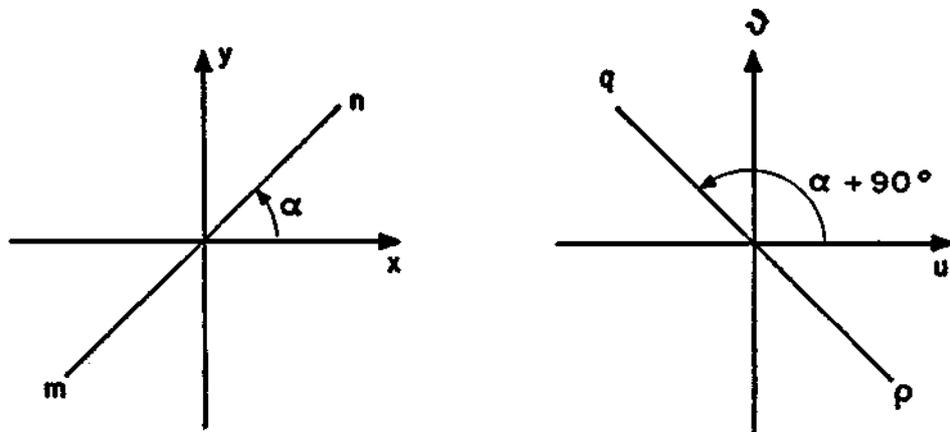
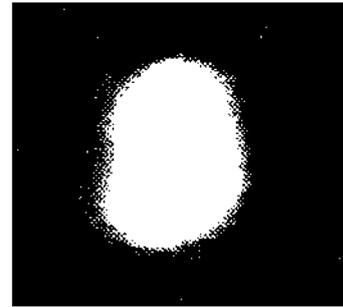


Fig. 4.2 - A Transformada de Fourier da imagem do tipo da reta paralela a $m-n$ no plano (x, y) é um corte da Transformada de Fourier Bidimensional ao longo da reta $p-q$ no plano (u, v) .



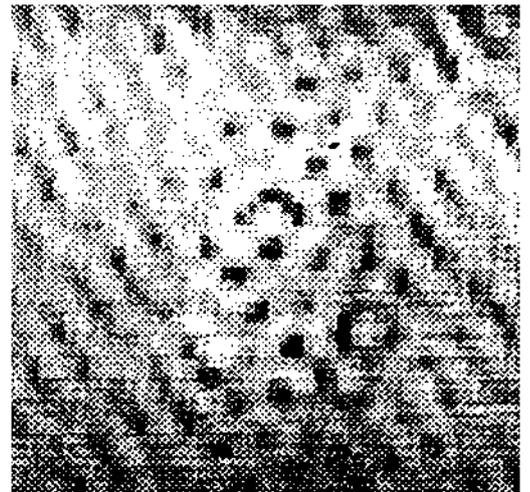
(a)



(b)

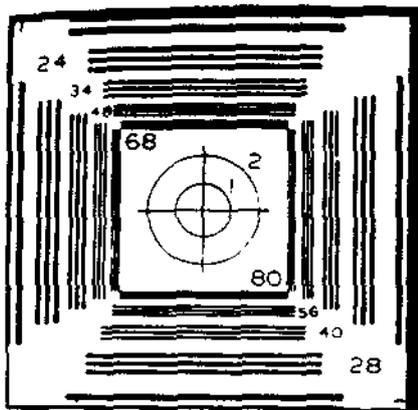


(c)

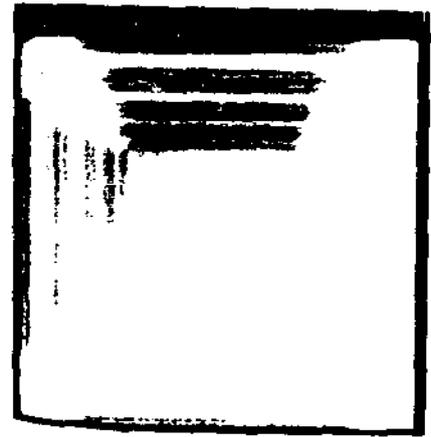


(d)

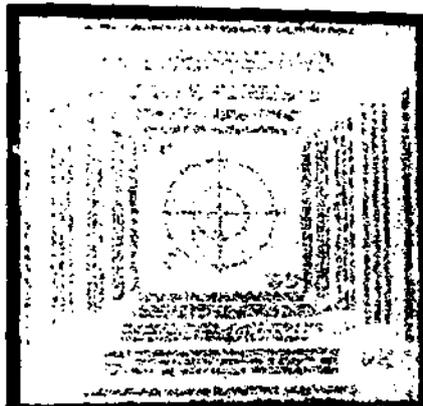
Fig. 4.3 - Restauração com filtro inverso: (a) imagem original; (b) imagem borrada; (c) restaruação obtida considerando uma vizinhança da origem do plano (u, v) que não inclui valores excessivamente pequenos de $H(u, v)$; (d) restauração obtida usando uma vizinhança maior onde essa condição não é válida.



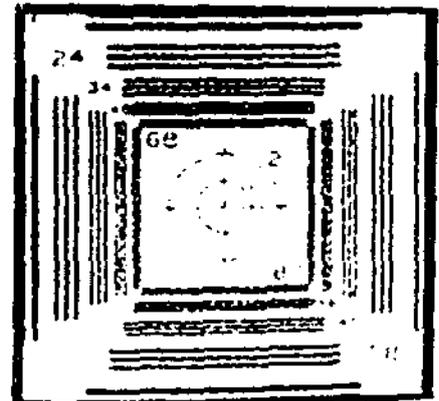
(a)



(b)



(c)



(d)

Fig. 4.4 - Restauração por mínimos quadrados com restrição: (a) imagem original; (b) imagem borrada com função de espalhamento do tipo gaussiano; (c) imagem restaurada pelo filtro inverso; (d) imagem restaurada por mínimos quadrados com restrição quadrática de igualdade.

CAPÍTULO 5

RECONSTRUÇÃO

O problema de reconstruir uma imagem por suas projeções tem atraído considerável atenção nos últimos anos, em particular devido à sua aplicação na área de tomografia computadorizada. Outras aplicações que têm despertado interesse na área médica incluem imageamento ultra-sônico e medicina nuclear. Desenvolve-se atividade também nos campos de radioastronomia, testes não-destrutivos de materiais, interferometria ótica, microscopia eletrônica e exploração geofísica. Além disso, há potencial para a aplicação das técnicas de reconstrução a várias outras áreas (Bates et alii, 1983).

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 sendo, em tomografia computadorizada, 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 do tecido naquela fatia. Um problema de reconstrução tridimensional é, normalmente, reduzido ao problema de reconstruir sucessivas fatias bidimensionais de um volume e empilhá-las. A Figura 5.1 ilustra a reconstrução de um corte do abdômem por tomografia computadorizada.

Várias técnicas têm sido propostas para abordar o problema de reconstrução de imagens. De modo geral, elas têm-se dividido em duas direções: a primeira delas utiliza as técnicas de transformadas da imagem e deriva seus métodos no domínio contínuo, discretizando os resultados no final para a implementação digital; a segunda direção (as chamadas técnicas algébricas) parte inicialmente de um modelo já discretizado, reduzindo o problema à redução de um grande sistema de equações lineares. Deve-se mencionar neste ponto que o problema de reconstruir um objeto bi ou tridimensional por suas projeções foi inicialmente abordado pelo matemático austríaco J. Randon em 1917, que introduziu a transformada que tem o seu nome. Todavia, a plena exploração das potencialidades dos algoritmos só foi possível com o computador digital, principalmente com o trabalho de G.N. Hounsfield, pelo qual foi ganhador do Prêmio Nobel de Medicina em 1979, juntamente com A.M. Carmack.

Deve-se observar também que o Prêmio Nobel de Química de 1982 foi também outorgado por trabalhos em reconstrução de imagens, envolvendo a reconstrução de vírus e macromoléculas por micrografia eletrônica.

O problema de reconstrução pode ser colocado em termos matemáticos da seguinte maneira. Seja $f(x, y)$ uma função definida no plano. Uma reta neste plano define um raio e a integral de $f(x, y)$ ao longo desse raio é uma integral de raio. O conjunto de integrais de raio segundo um dado ângulo forma uma projeção. A Figura 5.2 ilustra a projeção $P_{\theta}(R)$ de $f(x, y)$ segundo o ângulo θ .

A reta que define o raio segundo o ângulo θ e a distância R da origem é dada pela equação:

$$x \cos \theta + y \sin \theta = R \quad (5.1)$$

e a integral de $f(x, y)$ ao longo dessa reta pode ser expressa por:

$$P_{\theta}(R) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - R) dx dy \quad (5.2)$$

A função de duas variáveis $P_{\theta}(R)$ é a chamada Transformada de Radon de $f(x, y)$.

Os primeiros tomógrafos comerciais utilizavam feixes paralelos de raios, conforme ilustra a Figura 5.3.

Com a necessidade de abreviar o tempo de aquisição dos dados, em decorrência da mobilidade dos órgãos do corpo humano, os tomógrafos modernos passaram a utilizar uma geometria de raios divergentes conforme ilustra a Figura 5.4.

De modo geral, é possível estender as técnicas matemáticas desenvolvidas no caso de raios paralelos para a situação de raios divergentes. Assim sendo, neste capítulo, as deduções matemáticas partirão da hipótese de feixes de raios paralelos.

5.1 - TEOREMA DO CORTE DE FOURIER

A ferramenta básica para a reconstrução de imagens utilizando diretamente a Transformada de Fourier já foi elaborada no capítulo sobre restauração de imagens, quando se discutiu a determinação da F.E.P. pela imagem de retas em várias dimensões. O teorema estabelece essencialmente que a Transformada de Fourier de uma imagem bidimensional pode ser obtida pelas Transformadas de Fourier unidimensionais das projeções da imagem em várias direções. Considere-se, por exemplo, a projeção de $f(x, y)$ segundo o eixo x , dada por:

$$P_0(y) = \int_{-\infty}^{+\infty} f(x, y) dx \quad (5.3)$$

Deve-se observar que a variável y na Equação 5.3 faz o papel de R na Equação 5.2.

Tomando-se a Transformada de Fourier de $P_0(y)$ obtêm-se:

$$S_0(v) = \int_{-\infty}^{+\infty} P_0(y) e^{-j2\pi vy} dy, \quad (5.4)$$

ou ainda

$$S_0(v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-j2\pi vy} dx dy = F(u, v) \Big|_{u=0}, \quad (5.5)$$

isto é, a Transformada de Fourier da projeção na direção do eixo x é o corte da Transformada de Fourier de $f(x, y)$ na direção do eixo vertical no plano complexo.

Esse resultado pode ser generalizado para a projeção segundo um ângulo θ (Figura 5.5), dada por:

$$P_{\theta}(s) = \int_{-\infty}^{+\infty} f(t, s) dt, \quad (5.6)$$

onde t e s são obtidos de x e y pela mudança de variável

$$\begin{bmatrix} t \\ s \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \quad (5.7)$$

Obtêm-se, portanto, a Transformada de Fourier $S_{\theta}(w)$ de $P_{\theta}(s)$:

$$S_{\theta}(w) = \int_{-\infty}^{+\infty} P_{\theta}(s) e^{-j2\pi ws} ds = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(t, s) e^{-j2\pi ws} ds dt \quad (5.8)$$

Fazendo a mudança de variáveis de (s, t) para (x, y) na integração dada pela Equação 5.8 obtêm-se:

$$S_{\theta}(w) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P(t, s) e^{-j2\pi w(-x \sin \theta + y \cos \theta)} dx dy \quad (5.9)$$

= $F(u, v)$ para

$$\begin{aligned} u &= -w \sin \theta = w \cos (\theta + 90^{\circ}), \\ u &= w \cos \theta = w \sin (\theta + 90^{\circ}). \end{aligned} \quad (5.10)$$

As Equações 5.10 representam as equações paramétricas da reta indicada na Figura 5.5 b que faz um ângulo de $(\theta + 90^{\circ})$ com o eixo u , resultando em:

$$S_{\theta}(w) = F(w, \theta + 90^{\circ}). \quad (5.11)$$

Assim, tomando projeções em várias direções e obtendo as respectivas Transformadas de Fourier Unidimensionais, têm-se cortes sucessivos da Transforma

da de Fourier Bidimensional da Imagem, $F(u, v)$. Com infinitas projeções, $F(u, v)$ estaria completamente determinada e a imagem seria obtida pela Transformada de Fourier Inversa dada por:

$$f(x, y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u, v) e^{j2\pi(ux + vy)} du dv. \quad (5.12)$$

Como sō é disponível um número finito N^2 de amostras da Transformada de Fourier e, além disso, supõe-se que a função $F(x, y)$ é definida numa região limitada (por exemplo: $-\frac{A}{2} \leq x, y \leq \frac{A}{2}$), resulta em (Rosenfeld and Kak, 1982):

$$f(x, y) \cong \frac{1}{A^2} \sum_{m=-N/2}^{N/2} \sum_{n=-N/2}^{N/2} F\left(\frac{m}{A}, \frac{n}{A}\right) \exp \left[j2\pi \left(\frac{m}{A} x + \frac{n}{A} y \right) \right] \quad (5.13)$$

$$\frac{-A}{2} \leq x \leq \frac{A}{2}, \quad \frac{-A}{2} \leq y \leq \frac{A}{2}$$

A expressão 5.13 pode ser implementada pela Transformada Rápida de Fourier se os N^2 coeficientes $F\left(\frac{m}{A}, \frac{n}{A}\right)$ forem disponíveis.

Resta contudo o problema de que as projeções unidimensionais conduzem a amostras da Transformada de Fourier Bidimensional dispostas radialmente, ao passo que para efetuar a Transformada de Fourier Inversa Discreta Bidimensional, para recuperar a imagem original, as amostras devem estar dispostas numa grade retangular, surgindo assim a necessidade de interpolação. Teoricamente o problema pode ser resolvido determinando os N^2 coeficientes necessários para calcular a Expressão 5.13 se um número igual de valores é conhecido nas linhas radiais (Crowter et alii, 1970). Todavia, este cálculo envolve a solução de um conjunto grande de equações simultâneas, o qual frequentemente conduz à instabilidade numérica. É mais frequente determinar os valores na grade retangular por vizinho mais próximo ou por uma interpolação que envolve uma combinação linear das componentes próximas. Ainda assim, resta o fato de que os coeficientes de frequência espacial mais alta

são mais esparsos, implicando um maior erro de interpolação. Tem-se assim uma degradação na imagem obtida devido aos erros nas altas frequências espaciais.

O resultado do Teorema do Corte de Fourier foi derivado originalmente por Bracewell (1956) em Radioastronomia e depois independentemente por De Rosier e Klug (1968) em Microscopia Eletrônica e Rowley (1969) em Holografia Ótica. Atualmente, a maior aplicação desse teorema ainda é em Radioastronomia, onde é frequente a obtenção de dados diretamente no domínio de Fourier. Essa mesma razão torna o algoritmo adaptado a outras aplicações, como em tomografia por ressonância magnética nuclear (Hinshaw and Lent, 1983) ou por transmissão de ultra-som. (Mueller et alii, 1979).

5.2 - ALGORITMO DE FILTRAGEM E RETROPROJEÇÃO

A técnica de reconstrução mais utilizada nos tomógrafos comerciais modernos consiste no algoritmo denominado filtragem-retroprojeção (ou convolução-retroprojeção). A derivação desse resultado pode ser feita da seguinte forma.

Expressando a Transformada de Fourier Inversa em coordenadas polares obtêm-se:

$$f(x, y) = \int_0^{2\pi} \int_0^{\infty} F(w, \theta) e^{j2\pi w (x \cos \theta + y \sin \theta)} w \, dw \, d\theta. \quad (5.14)$$

Decompondo a integral de 0 a 2π e utilizando a relação:

$$F(w, \theta + \pi) = F(-w, \theta) \quad (5.15)$$

a equação 5.14 pode ser escrita como (Rosenfeld and Kak, 1982):

$$f(x, y) = \int_0^{\pi} \left[\int_{-\infty}^{+\infty} F(w, \theta) [w] e^{j2\pi w t} dw \right] d\theta. \quad (5.16)$$

Pelo Teorema do Corte de Fourier, $F(w, \theta) = S_{\theta - \pi/2}(w)$ (usando a notação da Figura 5.5), resultando em:

$$f(x, y) = \int_{-\pi/2}^{+\pi/2} \left[\int_{-\infty}^{+\infty} S_{\theta}(w) |w| e^{j2\pi w t} dw \right] d\theta, \quad (5.17)$$

onde t é expresso pela Equação 5.7.

A integral interior da Expressão 5.17 representa a Transformada de Fourier Inversa da função $S_{\theta}(w) |w|$, ou seja, é uma projeção "filtrada" pela função $|w|$. Denotando a função obtida por $Q_{\theta}(t)$ obtém-se:

$$Q_{\theta}(t) = \int_{-\infty}^{+\infty} S_{\theta}(w) |w| e^{j2\pi w t} dw \quad (5.18)$$

e a imagem representada por $f(x, y)$ é obtida "retroprojetando" as contribuições de todas as projeções filtradas pela integração dada por:

$$f(x, y) = \int_{-\pi/2}^{+\pi/2} Q_{\theta}(x \cos \theta + y \sin \theta) d\theta. \quad (5.19)$$

Assim, uma possível técnica de reconstrução seria calcular as Transformadas de Fourier $S_{\theta}(w)$ das projeções, multiplicá-las por $|w|$ e retroprojetá-las sobre a imagem. Uma alternativa seria utilizar o equivalente no espaço da multiplicação no domínio da frequência, dada pela operação de convolução, isto é:

$$Q_{\theta}(t) = P_{\theta}(t) * F^{-1}(|w|), \quad (5.20)$$

onde $*$ denota convolução.

Como se pode expressar:

$$|W| = (jw) \cdot (-j \operatorname{sgn} w), \quad (5.21)$$

onde a função $\operatorname{sgn} w$ é dada por:

$$\begin{aligned} \operatorname{sgn} w &= +1 & w > 0 \\ &= -1 & w < 0 \\ &= 0 & w = 0 \end{aligned} \quad (5.22)$$

tem-se que $Q_{\theta}(t)$ é dada pela Transformada de Hilbert da derivada de $P_{\theta}(t)$. Como $S_{\theta}(w)$ é limitada na prática a uma frequência, diga-se, W , deve-se trabalhar com uma versão limitada em faixa de $|w|$. O simples truncamento de $|w|$ pela multiplicação por uma janela do tipo pulso implica que $Q_{\theta}(t)$ seja dado pela Transformada de Hilbert da derivada de uma função do tipo $\frac{\operatorname{sent} t}{t}$ que tende a se tornar bastante oscilatória (Makovsky, 1976). Assim, têm sido utilizadas outras janelas com decaimento mais suave, como a de Hamming, (1977) por exemplo.

Além disso, as Equações 5.18 e 5.19 devem ser convenientemente discretizadas para implementação digital, resultando em (Rosenfeld and Kak, 1982):

$$Q_{\theta} \left(\frac{k}{2w} \right) \cong \frac{2W}{N} \sum_{n=-N/2}^{+N/2} S_{\theta} \left(m \frac{2W}{N} \right) \left[m \frac{2W}{N} \right] \cdot G \left(m \frac{2W}{N} \right) e^{j2\pi \left(\frac{mk}{N} \right)} \quad (5.23)$$

$$\tilde{f}(x, y) \cong \frac{\pi}{K} \sum_{i=1}^K Q_{\theta_i} (x \cos \theta_i + y \operatorname{sen} \theta_i), \quad (5.24)$$

onde $G \left(m \frac{2W}{N} \right)$ representa a janela utilizada e K é o número de ângulos em que

foram tomadas as projeções. A Equação 5.23 pode ser implementada eficientemente pela Transformada Rápida de Fourier.

O algoritmo de filtragem e retroprojeção pode ser interpretado a partir das Equações 5.23 e 5.24 ou 5.18 e 5.19. Para cada valor de θ_i , todos os pontos (x, y) na reta $t = x \cos \theta_i + y \sin \theta_i$ terão a mesma contribuição da projeção $Q_{\theta_i}(x \cos \theta_i + y \sin \theta_i)$. Tem-se então a idéia de retroprojeção. Por outro lado, se as projeções $P_{\theta}(t)$ não forem filtradas pela operação do tipo Transformada de Hilbert da derivada da projeção, pode-se mostrar que a reconstrução de uma imagem do tipo função delta de Dirac $\delta(x, y)$ ficará barrada, com uma função de espalhamento pontual do tipo $\frac{1}{r}$, onde r é a distância à origem (Herman, 1980; Rosenfeld and Kak, 1982).

5.3 - TÉCNICAS ALGÉBRICAS

As técnicas algébricas de reconstrução partem de um modelo discretizado, em que a imagem bidimensional a ser reconstruída é representada por um conjunto de $N \times N$ elementos de imagem. O valor observado no detector devido a uma projeção é o resultado da combinação linear dos valores dos elementos de imagem atravessados pela radiação, onde os pesos podem representar o tamanho da interseção do raio com a célula correspondente. A Figura 5.6 ilustra o esquema geométrico dos métodos algébricos de reconstrução. 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 elementos de imagem ser atravessada por ele.

5.3.1 - O Método de Kaczmarz

Uma das primeiras técnicas algébricas propostas para resolver o problema de reconstrução foi o algoritmo denominado ART ("Algebraic Reconstruction Technique") por Gordon et alii (1970). Este algoritmo foi posteriormente reconhecido como equivalente ao método de Kaczmarz para a resolução de um sistema de equações lineares (Kaczmarz, 1937). Esse método pode ser descrito do seguinte modo: as projeções obtidas geram um sistema de equações lineares dado por:

$$\sum_{j=1}^{N^2} w_{ij} f_j = p_i \quad i = 1, \dots, M, \quad (5.25)$$

onde M é o número total de raios projetados e N^2 é o número de células de imagem a serem reconstruídas. O método consiste num processo iterativo de solução do sistema linear através do seguinte passo:

$$f^{(j)} = (j-1) - \frac{\left(f^{(j-1)} \right)^T w_j - p_j}{\|w_j\|^2} w_j, \quad (5.26)$$

onde $f^{(j)}$ representa um vetor ($N^2 \times 1$) de estimações dos N^2 coeficientes de atenuação das células da imagem no $j^{\text{ésimo}}$ passo, e w_j é o vetor de pesos da $j^{\text{ésima}}$ linha do sistema de equações dado pela Equação 5.25. Estas equações dependem em geral da geometria do problema. O processo iterativo é iniciado com um vetor $f^{(0)}$ arbitrário. Pode ser mostrado (Rosenfeld and Kak, 1982) que a Equação 5.26 permite obter a $j^{\text{ésima}}$ estimativa de $f(f^{(j)})$ pela projeção de $f^{(j-1)}$ no hiperplano correspondente à $j^{\text{ésima}}$ equação do sistema dado pela Equação 5.25. O processo é cíclico, isto é, após utilizar a última equação volta-se à primeira e assim por diante. Quando existe uma única solução para o sistema de equações, Tanabe (1971) demonstrou que a sequência de projeções converge para a solução do sistema linear, conforme é ilustrado na Figura 5.7 para o caso simplificado de duas incógnitas. Já no caso de um sistema sobredeterminado de equações sujeitas à adição de ruído, as interseções dos hiperplanos não definem um único ponto e a sequência oscila ao redor do conjunto de interseções. No caso de um sistema subdeterminado há infinitas soluções e o processo converge para uma solução que minimiza a distância dessa solução ao vetor inicial escolhido.

Uma outra interpretação do método de Kaczmarz pode ser feita notando que $f^{(j-1)} \cdot w_j = q_j$ é a observação que seria obtida baseada nos va

lores dos elementos de imagem expressos pelo vetor $f_1^{(j-1)}$ da $(j^{\text{ésima}}-1)$ iteração. Assim sendo, o incremento $f_1^{(j)} - f_1^{(j-1)}$ na $j^{\text{ésima}}$ componente do vetor de estimativa \bar{e} é dado por:

$$f_1^{(j)} - f_1^{(j-1)} = \frac{(p_j - q_j) w_{j1}}{\sum_{k=1}^{N^2} w_{jk}^2} w_{j1}, \quad (5.27)$$

ou seja, subtrai-se a medida estimada na $(j^{\text{ésima}}-1)$ iteração da medida real, normaliza-se por $\sum_{k=1}^{N^2} w_{jk}^2$ e atribui-se esse incremento normalizado a cada $c_{\bar{e}}$ lula proporcionalmente ao peso w_{j1} da célula.

Além da eficiência computacional para trabalhar com grandes sistemas lineares esparsos, o método de Kaczmarz permite incorporar informação a priori sobre a imagem. A não-negatividade dos valores dos elementos de imagem pode ser conseguida tornando nulas as componentes negativas obtidas no decorrer do processo.

A dificuldade de implementar o método das projeções reside no cálculo e no armazenamento do grande número de coeficientes w_{ij} do sistema linear. Este problema pode ser abreviado considerando os coeficientes como uma função (que pode ser implementada por uma tabela) da distância entre o raio de projeção e o centro da célula; tal distância pode ser calculada em tempo de execução. Outra possibilidade consiste em substituir os coeficientes por 1 ou 0, conforme o raio atravessasse a célula ou não.

Essa última possibilidade leva a uma variação do método de Kaczmarz dada por:

$$f_1^{(j)} - f_1^{(j-1)} = \frac{p_j - q_j}{N_j}, \quad (5.28)$$

onde N_j é o número de células atravessadas pelo j -ésimo raio. Assim, o incremento $(p_j - q_j)$ é dividido igualmente entre todas as células atravessadas pelo raio.

Várias outras modificações do método das projeções têm sido propostas. Uma delas, sugerida por Gilbert (1972) realiza a atualização nos valores das densidades das células apenas no final de um ciclo completo de iterações. Outra possibilidade consiste na utilização dos chamados métodos de relaxação (Herman and Lent, 1976 a) em que, ao invés de projetar diretamente o vetor sobre o hiperplano, a sobre-relaxação (ultrapassar o hiperplano até a imagem especular) ou sub-relaxação (permanecer do mesmo lado do hiperplano) são procedimentos utilizados.

5.3.2 - Utilização de Desigualdades Lineares

É possível resolver o problema de reconstrução substituindo o sistema linear de igualdades dado pela Equação 5.25 por um sistema de desigualdades, por exemplo do tipo:

$$p_i - e_i \leq \sum_{j=1}^{N^2} w_{ij} f_j \leq p_i + e_i \quad i = 1, 2, \dots, M, \quad (5.29)$$

onde $e_i > 0$ representam tolerâncias que refletem a incerteza nas medidas atribuída ao ruído. Devido à magnitude e esparsidade do sistema de desigualdades, é usual utilizar os métodos iterativos (Censor, 1981).

Um método possível é o de relaxação para desigualdades de Agmon (1954) e Motzkin e Schoenberg (1954). Nesse método o fato de as desigualdades aparecerem em pares não é levado em consideração. Multiplicando o lado esquerdo das Desigualdades 5.29 por (-1) obtém-se um sistema com o dobro do número de desigualdades, o qual pode ser expresso por:

$$\langle \underline{w}_i, \underline{f} \rangle \leq d_i \quad i = 1, 2, \dots, 2M \quad (5.30)$$

e o processo iterativo é dado por:

$$\underline{f}^{(0)} \text{ arbitrário,} \tag{5.31}$$

$$\underline{f}^{(j+1)} = \underline{f}^{(j)} + s_j \underline{w}_i, \tag{5.32}$$

$$\text{onde } s_j = \min \left\{ 0, \lambda_j \frac{d_i - \langle \underline{w}_i, \underline{f}^{(j)} \rangle}{\|\underline{w}_i\|^2} \right\}. \tag{5.33}$$

A sequência $\{\lambda_j\}_{j=0}^{\infty}$ representa parâmetros de relaxação restritos ao intervalo $\delta_1 \leq \lambda_j \leq 2 - \delta_2$, $\delta_1, \delta_2 > 0$ e garante a convergência de algoritmo. A sequência de iterações percorre ciclicamente as equações de projeções.

Outro tipo natural de restrição do tipo desigualdade linear que se pode incorporar envolve limites inferiores e superiores nos valores estimados. Frequentemente um limite inferior natural é zero, como por exemplo no caso de a componente f_i representar o coeficiente de absorção de raios X da i -ésima célula. Um limite superior pode representar, por exemplo, o maior coeficiente de absorção do corpo humano dado pelo osso compacto. O método procura uma solução de mínima norma que satisfaça as desigualdades devidas à imprecisão nas medidas e os limites inferiores e superiores, isto é, procura-se um vetor \underline{f} que minimize $1/2 \|\underline{f}\|^2$ sujeito às restrições $\gamma_i \leq \langle \underline{w}_i, \underline{f} \rangle \leq \delta_i$ para $i = 1, 2, \dots, M$ e $\underline{w}_i \leq f_i \leq v_i$ para $i = 1, 2, \dots, N^2$. A descrição do processo iterativo para a solução desse problema de programação quadrática poderá ser encontrada no trabalho original de Herman e Lent (1978) ou Censor (1981).

É conveniente notar que o problema de resolver sistemas de desigualdades lineares ocorre também nos algoritmos de treinamento em classificação de padrões (Duda and Hart, 1973), conforme será mostrado no Capítulo 10. Além disso, o problema de restauração de imagens pode também ser colocado em termos semelhantes (Mac Adam, 1970, Mascarenhas et alii, 1978).

5.3.3 - Algoritmos de Mínimos Quadrados - Métodos de Suavização - Métodos Bayesianos

Do mesmo modo que o problema de restauração de imagens, o problema de reconstrução, quando discretizado, leva à solução de um grande sistema linear de equações. É natural, portanto, que métodos semelhantes aqueles desenvolvidos para restauração sejam aplicadas em reconstrução.

Quando se procura uma solução de mínimos quadrados sem incorporar informação a priori sobre a imagem, os problemas de mal condicionamento podem surgir (como ocorrem também nos métodos de transformadas), particularmente quando se dispõe de observações em ângulos limitados, o que tende a provocar dependência linear das equações de projeção. Dois métodos inicialmente propostos para reconstrução (Gilbert, 1972 e Goiten, 1972) foram posteriormente identificados (Larshminarayanan and Lent, 1972 e Herman and Lent, 1976 b) como métodos de mínimos quadrados. Kashyap e Mittal (1975) também propuseram dois métodos, um dos quais foi reinterpretado como de mínimos quadrados e o outro como uma formulação bayesiana (Mascarenhas and Fernandes, 1980). Essas reinterpretações levam a formulações recursivas que, como os métodos iterativos de solução de equações lineares, adaptam-se à maneira usualmente sequencial com que são tomados os dados em reconstrução. A recursividade pode ser feita também em blocos (Mascarenhas and Fernandes, 1980; Eggermont et alii, 1981), atualizando a estimação após as medidas correspondentes a um ângulo de projeção.

As técnicas de suavização e regularização também têm sido empregadas em reconstrução de imagens, na forma de algoritmos iterativos capazes de ser utilizados em minicomputadores com pequena memória principal (Herman et alii, 1980).

Deve-se observar que Herman e Lent (1976 b) propuseram uma formulação unificada determinística dos métodos que minimizam formas quadráticas sujeitas a restrições quadráticas de igualdades (os quais englobam os métodos de suavização ou ainda, por uma reinterpretação adequada, os estimadores bayesianos lineares) do seguinte modo:

$$\text{Min } \|\underline{D}^{-1} \underline{f}\|^2 \text{ tal que} \quad (5.34)$$

$$a(\underline{p} - \underline{Wf})^T \underline{A}(\underline{p} - \underline{Wf}) + (\underline{f} - \underline{\bar{f}})^T (b \underline{B} + c \underline{C}^{-1}) (\underline{f} - \underline{\bar{f}}) \quad (5.35)$$

onde \underline{f} é o vetor de densidades das células da imagem; \underline{p} é o vetor de medidas das projeções; \underline{W} é a matriz de projeção; \underline{A} é uma matriz positiva semidefinida; \underline{B} , \underline{C} e \underline{D} são matrizes positivas definidas; a , b e c são números reais não-negativos; e $\underline{\bar{f}}$ é um vetor especificado.

5.3.4 - Otimização da Entropia

A abordagem de máxima entropia procura obter uma solução para o problema de reconstrução com o mínimo conteúdo de informação compatível com os dados obtidos. Uma típica formulação do problema é dado por (Gordon et alii, 1970)

$$\text{Max } \left(- \sum_{j=1}^{N^2} f_j \ln f_j \right) \quad (5.36)$$

tal que

$$\langle \underline{w}_i, \underline{f} \rangle = p_i \quad i = 1, 2, \dots, M \quad (5.37)$$

e

$$\underline{f} \geq \underline{0}. \quad (5.38)$$

O algoritmo que segue foi proposto por Gordon et alii (1970) e sua convergência foi provada por Lent (1977) sob determinadas condições:

$$\underline{f}^{(0)} = e^{-1} \cdot \underline{1} \quad (5.39)$$

$$\underline{f}_j^{(k+1)} = \left[\frac{p_i}{\langle \underline{w}_i, \underline{f}^{(j)} \rangle} \right]^{\lambda^k w_i^2} \underline{f}_j^{(k)} \quad j = 1, 2, \dots, N^2, \quad (5.40)$$

onde e é a base dos logaritmos naturais, $\underline{1}$ é um vetor com todas as componentes unitárias e λ_k são parâmetros de relaxação, tal que $0 < \lambda_k \leq 1$. As equações são tomadas ciclicamente, i.e., $i = k(\text{mod } M) + 1$. A condição necessária para a convergência do algoritmo é que o conjunto expresso pelas Restrições 5.37 e 5.38 seja não-vazio. Não é conhecido o comportamento do algoritmo quando aquela condição não é satisfeita.

Os métodos baseados na maximização da entropia têm sido empregados na reconstrução de imagens, tanto em técnicas algébricas (Minerbo, 1979) como em técnicas que envolvem transformadas (Wernecke, 1977). O critério de máxima entropia parece ser conveniente nas situações com deficiência de dados obtidos, como por exemplo na redução da dosagem de raios X em pacientes.

5.4 - COMPARAÇÃO ENTRE MÉTODOS - PROBLEMAS ABERTOS

O tomógrafo computadorizado desenvolvido inicialmente por Hounsfield utilizava uma técnica algébrica iterativa de reconstrução. A maior rapidez dos processos de reconstrução por convolução e retroprojeção levou os tomógrafos modernos a utilizarem quase que exclusivamente este último método. Todavia, existem várias situações (discutidas em detalhe por Censor, 1983) onde os métodos algébricos podem oferecer um desempenho superior. Entre elas podem-se mencionar a possibilidade de incorporar conhecimento a priori, a flexibilidade de trabalhar com problemas onde as projeções não são uniformemente espaçadas ou com modelos não-lineares, etc.

Dentre os problemas abordados na atualidade, destacam-se:

- 1) Estudo de tomografia computadorizada por emissão, onde a necessidade de estimar simultaneamente a atenuação e a distribuição da fonte torna o problema não-linear.
- 2) Reconstrução por dados incompletos; se a construção de um objeto é feita normalmente para $-b \leq R \leq b$ e $0 \leq \theta \leq \pi$, Figura 5.2, podem-se considerar 3 casos:

- a) projeções truncadas - dados disponíveis para $-b < -a \leq R \leq a < b$, $0 \leq \theta < \pi$;
 - b) projeções "ocas" - dados disponíveis para $-b \leq R \leq a$, $a \leq R \leq b$, $0 \leq \theta \leq \pi$, e
 - c) faixa limitada de visão - dados disponíveis para $-b \leq R \leq b$, $0 < \alpha \leq \theta \leq \beta < \pi$. Para uma discussão pormenorizada desses problemas, o leitor pode consultar Altschuller et alii (1981) ou Louis e Natterer (1983);
- 3) Reconstrução tridimensional, motivada principalmente pela necessidade de reconstruir órgãos móveis como o coração humano, por exemplo, onde a reconstrução sequencial por fatias se torna extremamente difícil. A reconstrução tridimensional apresenta também o problema de fazer a exibição desses órgãos reconstruídos, o que envolve a solução de problemas de detecção de contornos de superfícies dos órgãos, bem como de problemas de gráficos por computador, como a eliminação de linhas e superfícies escondidas, visão por perspectivas, etc. (Udupa, 1983).

REFERENCIAS BIBLIOGRÁFICAS

- AGMON, S. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 6:382-392., 1954.
- ALTSCHULER, M.D.; CENSOR, Y.; HERMAN, G.T.; LENT, A.; LEWITT, R.M.; SRIHARI, S.N.; TUY, H.; UDUPA, J.K. Mathematical aspects of image reconstructions from projections. In: Kanal, L.N.; Rosenfeld, A.; ed. *Progress in pattern recognition*. Amsterdam, North Holland, 1981, p. 323-376.
- BATES, R.H.T.; GARDEN, K.L.; PETERS, T.M. Overview of computerized tomography with emphasis on future developments. *Proceedings of the IEEE*, 71(5):356-372, Mar. 1983.
- BRACEWELL, R.N. Strip integration in radioastronomy. *Australian Journal of Physics*, 9:198-217, 1956.
- CENSOR, Y. Row-action methods for huge and sparse systems and their applications. *SIAM Review*, 23:444-464, 1981.
- Finite series expansion reconstruction methods. *Proceedings of the IEEE*, 71(3):409-419, 1983. Special Issue on Computerized Tomography.
- CROWTER, R.A.; DE ROSIER, D.J.; KLUG, A. The reconstruction of a three-dimensional structure from projections and its applications to electron microscopy. *Proceedings of the Royal Society of London. Ser A*, 317:319-340, 1970.
- DE ROSIER, D.J.; KLUG, A. Reconstruction of three-dimensional structures from electrons micrographs. *Nature*, 217:130-134, 1968.
- DUDA, R.O.; HART, P. *Pattern classification and scene analysis*. New York, NY., John Wiley, 1973.
- EGGERMONT, P.P.B.; HERMAN, G.T.; LENT, A. Iterative algorithms for large partitioned linear systems with application to image reconstruction. *Linear Algebra its Applications*, 40:37-67, 1981.
- GILBERT, P. Iterative methods for the reconstruction of three-dimensional objects from their projections. *Journal Theoretical Biology*, 36:105-117, 1972.
- GOITEN, M. Three-dimensional reconstructions from a series of two-dimensional projections. *Nuclear Instruments and Methods*, 101:509-518, 1972.

- GORDON, R.; BENDER, R.; HERMAN, G.T. Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *Journal Theoretical Biology*, 29:471-481, 1970.
- HAMMING, R.W. *Digital filters*. Englewood Cliffs, NJ. Prentice Hall, 1977.
- HERMAN, G.T. *Image reconstruction from projections: the fundamentals of computerized tomography*. New York, NY, Academic, 1980.
- HERMAN, G.T.; LENT, A. A family of iterative quadratic optimization algorithms for pairs of inequalities, with application in diagnostic radiology. *Mathematical Programming Study*, 9(8):15-29, 1978.
- Iterative reconstruction algorithms. *Comput Biol. and Med.*, 6:273-294, 1976a.
- Quadratic Optimization for Image Reconstruction I. *Computer Graphics Image Processing*, 5(SF):319-332, 1975b.
- HERMAN, G.T.; LENT, A.; HURWITZ, H. A storage-efficient algorithm for finding the regularized solution of a large inconsistent system of equations. *Journal of the Institute of the Mathematics and Applications*, 25:361-366, 1980.
- HINSHAW, W.S.; LENT, A. An introduction to NMR imaging: from the Bloch equation to the imaging equation, *Proceedings of the IEEE*, 71(3):338-356, 1983.
- KACZMARZ, S. Angenähert auflosung van systemen linearer gleichungen. *Bulletin of the Acad. Polon. Sci. Lett., A.*, 35:355-357, 1937.
- KASHYAP, R.L.; MITTAL, M.C. Picture reconstruction from projections. *IEEE Transactions on Computers*, CAS-24(9):915-923, Sept. 1975.
- LAKSMINARAYAMAN, A.V.; LENT, A. Methods of least squares and SIRT in reconstruction. *Journal of Theoretical Biology*, 76:267-295, 1979.
- LENT, A. Maximum entropy and multiplicative ART. In: SHAW, R., ed *Image analysis and evaluation*. Washington, DC, Society of Photographic Scientists and Engineers, 1977.
- LOUIS, A.K.; NATTERER, F. Mathematical problems of computerized tomography. *Proceedings of the IEEE*, 71(3):379-389, Mar. 1983.

- MAC ADAM, D.P. Digital image restoration by constrained deconvolution. *Journal of the Optical Society of America*, 60:1617-1627, 1970.
- MAKOVSKI, A. *Medical imaging system*; notas de aula. Stanford University. Stanford, 1976.
- MASCARENHAS, N.D.A.; FERNANDES, L.F.V. New methods for picture reconstruction: recursive and causal techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2(4):369-376, 1980.
- MASCARENHAS, N.D.A.; FERNANDES, L.F.V.; SOUZA, M.L.C. Restauração de imagens utilizando algoritmos de treinamento do tipo "perceptron". In: CONGRESSO INTERNACIONAL DE PROCESSAMENTO DE SINAIS, Rio de Janeiro, 1978. *Anais*. Rio de Janeiro, UFRJ, 1978, p.15-25.
- MINERBO, G. Ment: A maximum entropy algorithm for reconstructing a source from projection Data. *Computer Graphics and Image Processing*, 10:48-68, 1979.
- MOTZKIN, T.S.; SCHOENBERG, I.J. The relaxation method for linear inequalities. *Canadian Journal of Mathematics*, 16:393-404, 1954.
- MUELLER, R.K.; KAVEH, M.; WADE, G. Reconstructive tomography and applications to ultrasonics. *Proceedings of the IEEE*, 67:567-587, 1979.
- ROSENFELD, A.; KAK, A.C. *Digital picture processing*. 2ª ed. New York, NY, Academic, 1982.
- ROWLEY, P.D. Quantitative interpretation of three-dimensional weakly refractive phase objects using holographic interferometry. *Journal Optical Society of America*, 59:1594-1498, 1969.
- TANABE, T. Projection method for solving a singular system of linear equations and its applications. *Numerische Mathematik*, 17:203-214, 1971.
- UDUPA, J.K. Display of 3D information in discrete 3D scenes produced by computerized tomography. *Proceedings of the IEEE*, 71(3):420-431, Mar. 1983.
- WERNECKE, S.J.; D'ADDARIO, L.R. Maximum entropy image reconstruction. *IEEE Transactions Computer*, C-26:351-364, 1977.

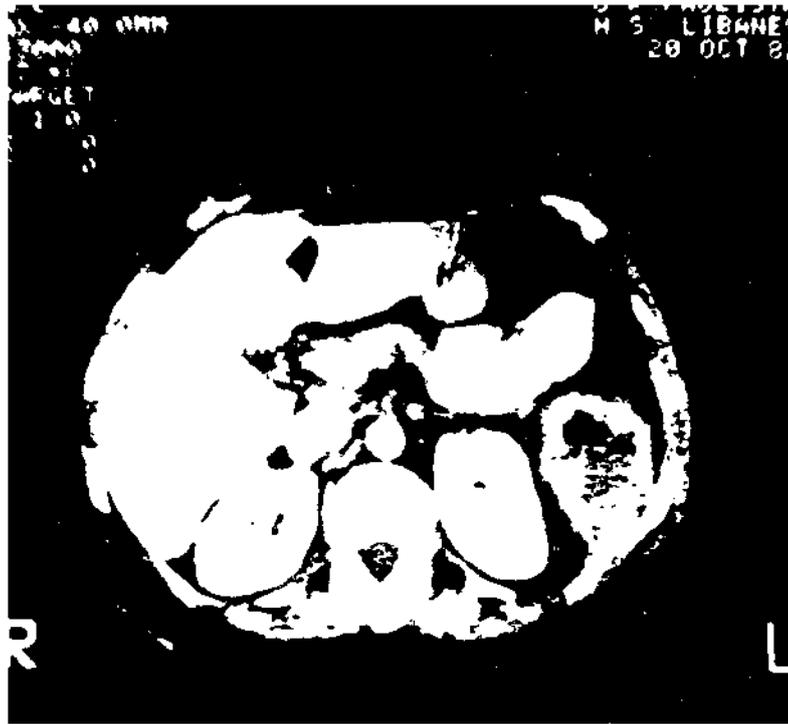


Fig. 5.1 - Reconstrução tomográfica por raios x - corte do abdômem.

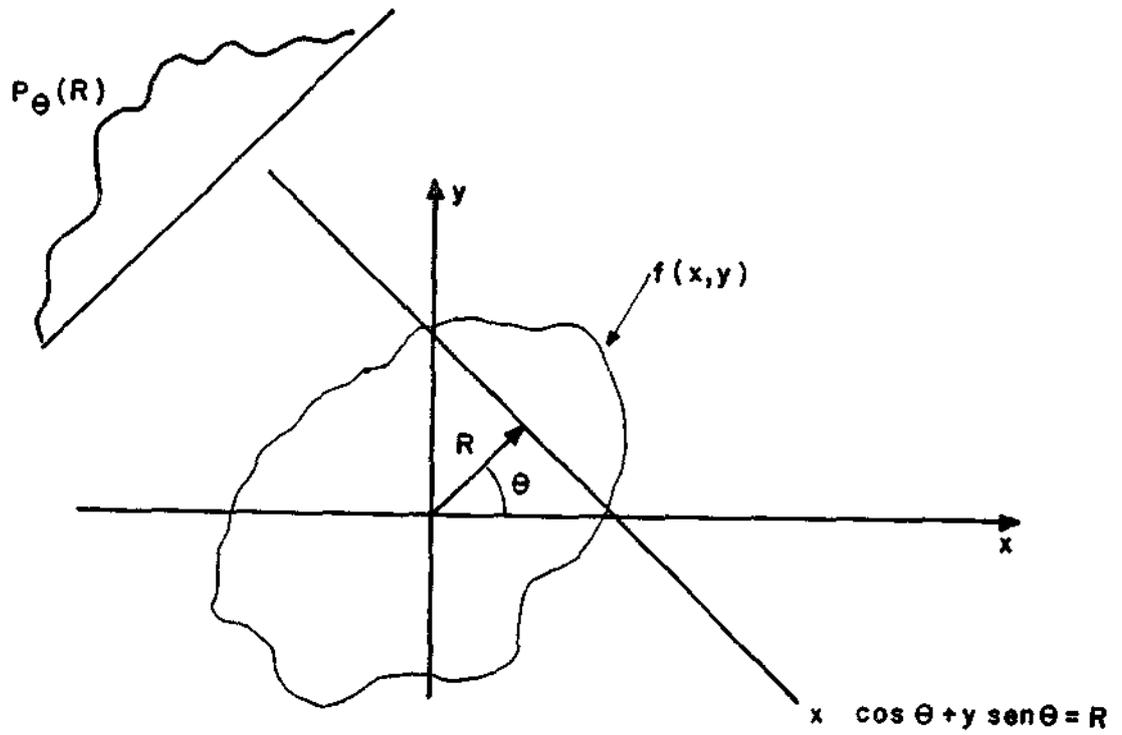


Fig. 5.2 - Projeção $P_{\theta}(R)$ de $f(x, y)$ segundo o ângulo θ .

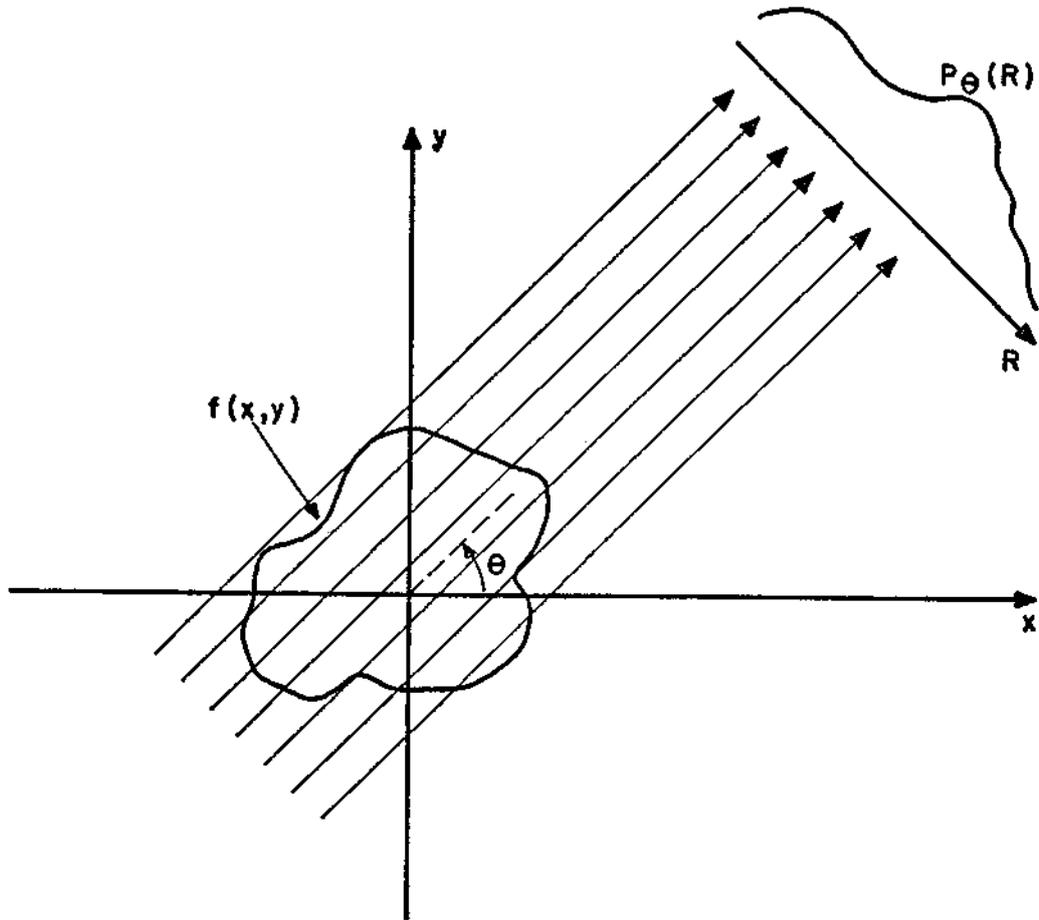


Fig. 5.3 - Projeção segundo raios paralelos.

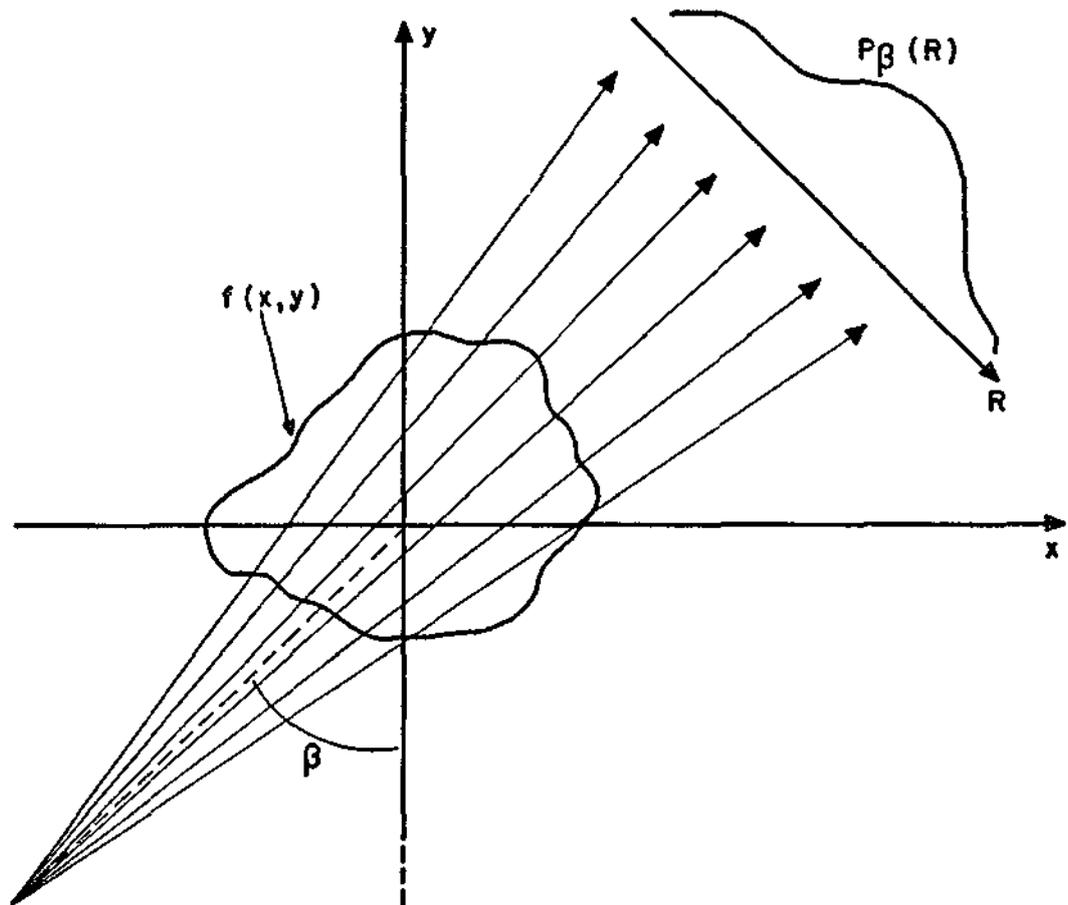


Fig. 5.4 - Projeção segundo raios divergentes.

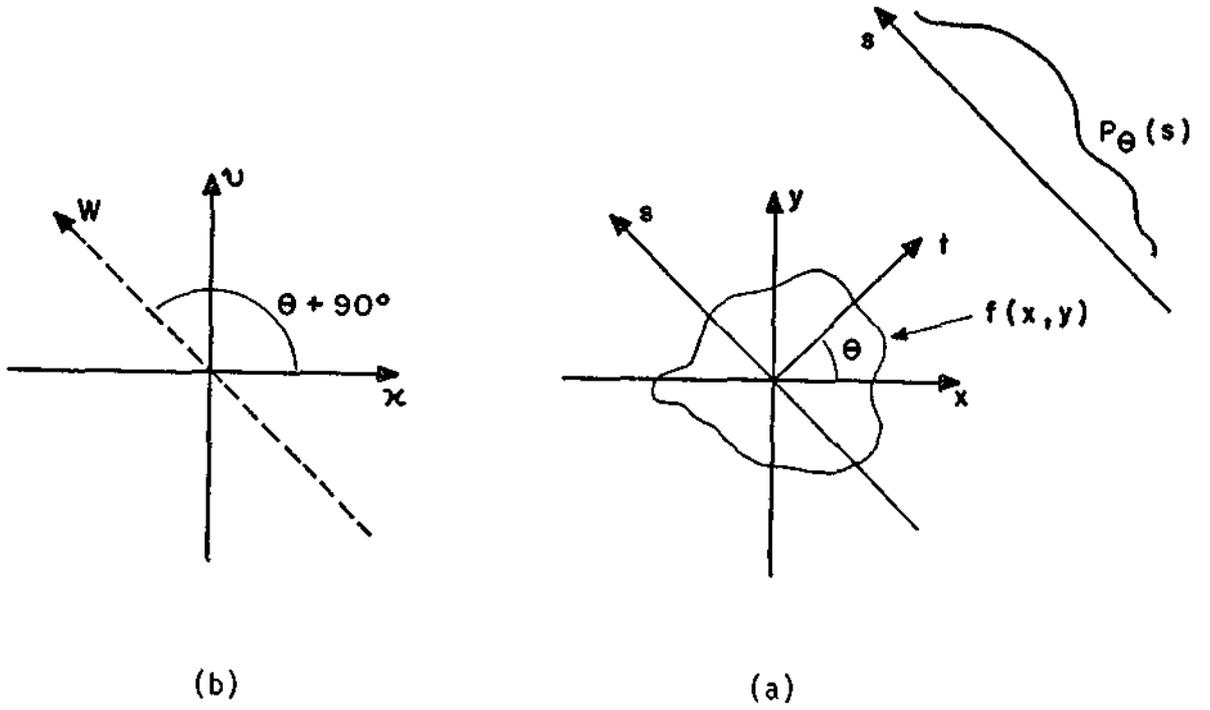


Fig. 5.5 - Transformada de Fourier de $P_{\theta}(s)$ Fig.(a): corte da Transformada de Fourier Bidimensional $F(u,v)$ na direção da reta indicada na Figura (b).

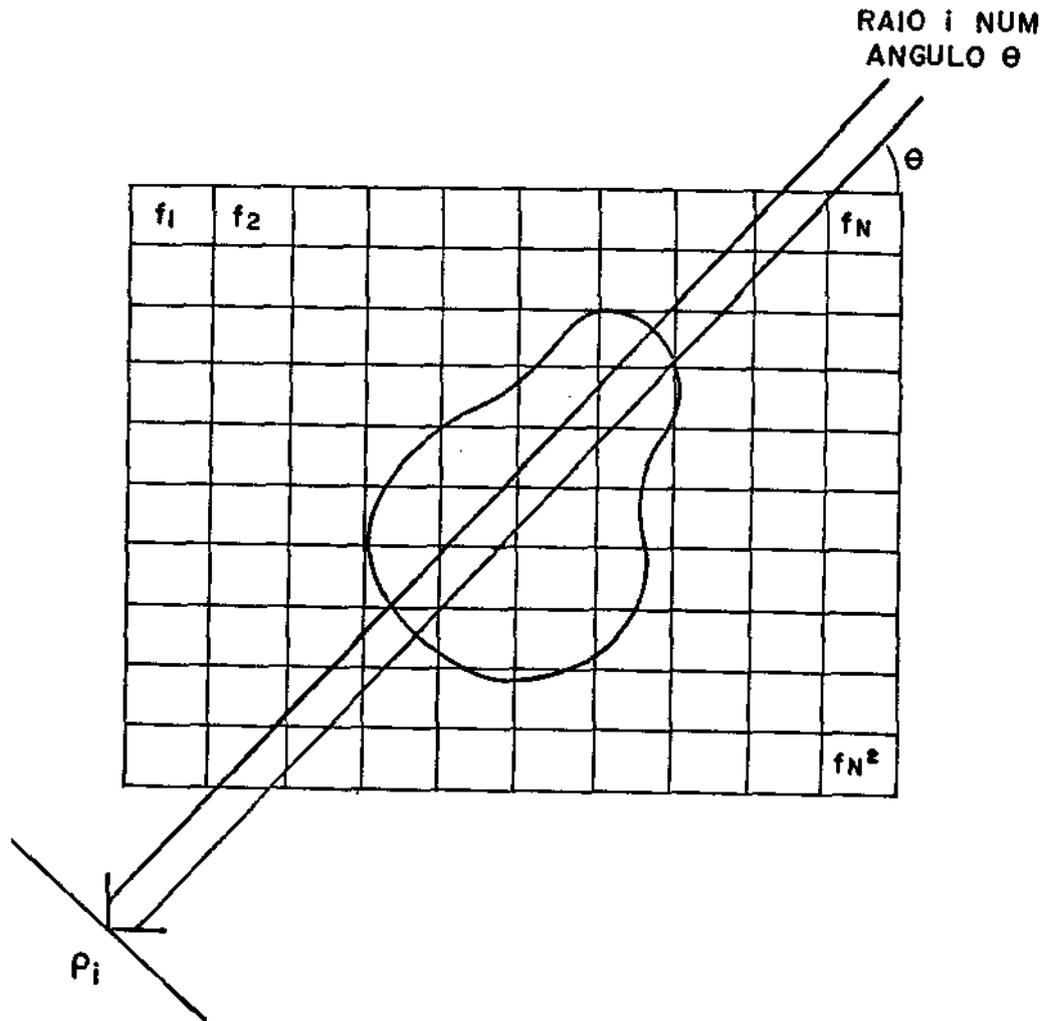


Fig. 5.6 - Esquema geométrico dos métodos algébricos de reconstrução.

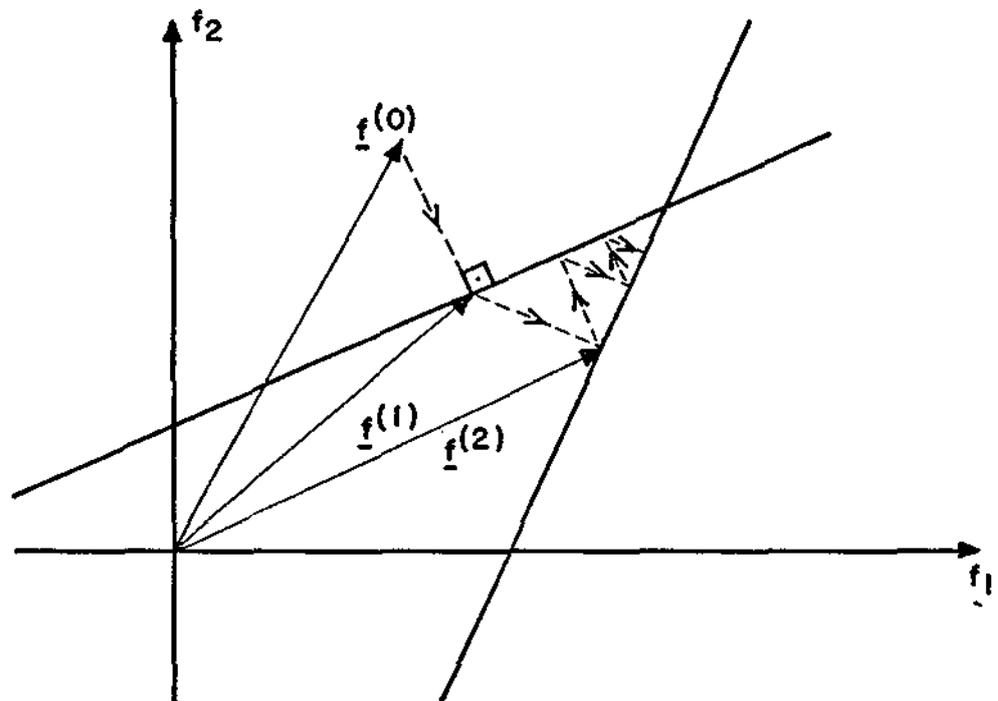


Fig. 5.7 - Interpretação geométrica do método de Kaczmarz - Solução única.

CAPÍTULO 6

CODIFICAÇÃO

O desenvolvimento das técnicas de comunicação digitais tem provocado grande interesse na codificação digital da informação pictórica. As maiores potencialidades residem na transmissão ponto-a-ponto mais do que na difusão de imagens de TV, onde já existe um enorme investimento em receptores domésticos analógicos. Assim sendo, desenvolvem-se esforços para digitalizar a transmissão de imagens em redes de distribuição de TV, em sistemas de teleconferência ou em transmissão de documentos por fac-símile. Além disso, abrem-se possibilidades de armazenamento eficiente de imagens biomédicas como radiografias, gráficos de engenharia ou arquivos de impressões digitais.

De um ponto de vista geral, o objetivo dos processos de codificação de imagens 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, necessárias para transmitir a imagem, ou o volume de memória para armazená-la.

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

A Modulação com Codificação por Pulsos (em inglês "Pulse Code Modulation" - PCM) é a forma mais simples de codificação de imagens. 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 ($M=2^B$), o que resulta numa codificação em B bits. Este sistema não elimina qualquer redundância nos elementos de 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 discussão sobre redução dos níveis de quantização apresentada no Capítulo 2 é válida também aqui.

6.1 - O PROBLEMA DE CODIFICAÇÃO DA FONTE DE SINAL - TEORIA DE RAZÃO DE DISTORÇÃO

A redundância estatística de uma imagem caracterizada como uma fonte de sinal pode ser expressa em termos da entropia da fonte em bits por imagem como:

$$H = - \sum_{i=1}^T p_i \log_2 p_i, \quad (6.1)$$

onde p_i é a probabilidade a priori da i ésima imagem e T é o número total das imagens que pode ser produzido pela fonte. Em princípio, uma codificação ótima exata poderia ser obtida pelo código de Huffman (a ser discutido adiante neste Capítulo), atribuindo um código mais curto às imagens que ocorrem mais frequentemente. Todavia, o número de imagens que pode ser produzido pela fonte é tão grande (por exemplo, se uma imagem contém 256 por 256 elementos, cada um quantizado com $256 = 2^8$ níveis, o número total de imagens possíveis é dado por $2^{8 \times 256 \times 256} = 2^{524588}$) o que torna completamente inviável esta abordagem do problema. Mesmo no caso de tentar uma codificação em blocos de modo a diminuir T , o problema torna-se viável apenas para pequenos blocos, envolvendo 2 ou 3 vizinhos para imagens monocromáticas e da ordem de 10 vizinhos para imagens binárias de fac-símile.

Se, por outro lado, for tolerada uma certa distorção média D na representação de uma classe de imagens, é natural perguntar-se qual o mínimo número médio de bits necessário para representar as imagens naquela classe. Essa resposta é fornecida pela teoria de razão de distorção, iniciada por Shannon (1949) e subsequentemente desenvolvida por inúmeros pesquisadores (Berger, 1971). Todavia, esta teoria fornece apenas os limites dos métodos de codificação, não especificando como projetá-los. Além disso, ela sofre as deficiências no modelamento estatístico da fonte de sinal (imagens), na dificuldade de cálculo da função de razão de distorção $R(D)$ (mínima taxa de bits necessária para transmissão ou armazenamento sob uma distorção média D) e no estabelecimento de um critério de distorção visual que, simultaneamente seja tratável do ponto de vista matemático e leve em conta as características complexas do sistema visual humano. Apesar dessas limitações, a teoria de razão de distorção oferece a base matemática

ca para o problema de compressão de dados. Assim sendo, será apresentada a seguir uma breve introdução às noções envolvidas, baseada na apresentação de Rosenfeld and Kak (1982).

Considere-se uma fonte de sinais que gere sucessivos e independentes símbolos $\{a_j\}$ que são identicamente distribuídos. Esses símbolos passam por um canal discreto e sem memória, isto é, cada símbolo a_j é processado independentemente, o que resulta na saída um símbolo b_k . O canal é caracterizado pelo conjunto de probabilidades condicionais $P(b_k/a_j)$. Seja P_C o conjunto dessas probabilidades condicionais. Assim,

$$P_C = \left\{ P(b_k/a_j), \begin{matrix} k = 1, 2, \dots, n \\ j = 1, 2, \dots, m \end{matrix} \right\}. \quad (6.2)$$

Suponha-se que, tendo sido enviado pelo canal o símbolo a_k e recebido o símbolo b_j , a penalidade (não-negativa) atribuída a tal fato seja dada por $\rho(a_j, b_k)$, que é uma medida de distorção. O valor médio da distorção é especificada por:

$$\begin{aligned} d(P_C) &= \sum_{j=1}^m \sum_{k=1}^n \rho(a_j, b_k) P(a_j, b_k) \\ &= \sum_{j=1}^m \sum_{k=1}^n \rho(a_j, b_k) P(a_j) P(b_k/a_j). \end{aligned} \quad (6.3)$$

Note-se que a distribuição de probabilidade $\{P(a_j) \ j=1, \dots, m\}$ caracteriza a fonte de sinal.

Um canal é denominado D-admissível se $d(P_C) \leq D$ para todas as possíveis distribuições $\{P(a_j), \ j=1, 2, \dots, m\}$. O conjunto de todas as distribuições condicionais de probabilidade P_C (que caracteriza o canal) D-admissíveis é denotado por:

$$P_D = \{P_C / d(P_C) \leq D\}. \quad (6.4)$$

Para uma fonte e um canal, a razão de transferência de informação do receptor ao transmissor é dada pela informação média mútua (Gallager, 1968) especificada por:

$$H(X;Y) = \sum_{j=1}^m \sum_{k=1}^n P(a_j) P(b_k/a_j) \log \frac{P(b_k/a_j)}{P(b_k)}. \quad (6.5)$$

Define-se então a função de razão de distorção $R(D)$ como:

$$R(D) = \min_{P_C \in P_D} H(X;Y), \quad (6.6)$$

para uma medida de distorção $\rho(a_j, b_k)$ escolhida. Assim sendo, $R(D)$ é a mínima razão de transferência de informação da fonte especificada ao receptor para que o receptor possa reproduzir os símbolos dessa fonte com uma distorção média D . $R(D)$ é medida em nats ou bits por símbolo.

Pode-se mostrar que $R(D)$ possui as seguintes propriedades (Berger, 1971):

- a) $R(D)$ não é definida para $D < 0$;
- b) existe um valor $D_{\max} \geq 0$ dado por

$$D_{\max} = \min_k \sum_{j=1}^m p(a_j) \rho(a_j, b_k) \quad (6.7)$$

tal que $R(D) = 0$ para $D \geq D_{\max}$;

- c) $R(D)$ é positiva, monotônica decrescente e contínua para $0 < D < D_{\max}$;
- d) $R(D)$ é convexa, isto é, para qualquer D' e D'' e qualquer λ tal que $0 \leq \lambda \leq 1$, tem-se que

$$R(\lambda D' + (1 - \lambda) D'') \leq \lambda R(D') + (1 - \lambda) R(D''). \quad (6.8)$$

Essas propriedades mostram que uma curva típica $R(D)$ deve ter a forma da Figura 6.1.

Deve-se notar que a existência do canal na discussão precedente é puramente formal. Assim sendo, no problema de quantização, a fonte de sinal pode ser identificada como os elementos contínuos de imagem a serem quantizadas e o canal como o amostrador e o quantizador. Observe-se que neste caso a função de razão de distorção deveria ser definida para fontes contínuas de símbolos.

Voltando ao exemplo anteriormente apresentado de uma imagem discretizada de 256×256 elementos quantizados com 8 bits, para a aplicação da teoria de razão de distorção, se cada possível imagem for associada a um símbolo a_j , além da necessidade de estipular a distribuição de probabilidade $P(a_j)$ para os $2^{8 \times 256 \times 256}$ possíveis símbolos da fonte, seria necessário também especificar os $2^{2 \times 8 \times 256 \times 256}$ possíveis valores de $\rho(a_j, b_k)$.

Uma abordagem mais tratável seria encarar os possíveis níveis de cinza dos elementos de imagem individuais como os símbolos $\{a_j\}$. Todavia, a medida de distorção $\rho(a_j, b_k)$ tenderia a ser excessivamente local do ponto de vista psicofísico pois, conforme foi examinado no Capítulo 2, um critério de distorção deve levar em conta o contexto espacial, já que a visão humana tolera maior distorção em regiões de grande detalhe, o oposto acontecendo com regiões de variação suave nos níveis de cinza.

6.2 - TÉCNICAS REVERSÍVEIS

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 visual é tolerada. Entre as técnicas reversíveis de exatas deve-se mencionar o código de Huffman (Huffman, 1952), que associa códigos mais curtos aos tons de cinza mais frequentes na imagem, minimizando assim a taxa de bits. Códigos estatísticos do tipo Huffman têm sido aplicados em situações onde não se deve ter perda de informação por exemplo, por razões legais como em imagens 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.

6.2.1 - Codificação de um Elemento de Imagem

É possível aplicar o código de Huffman diretamente aos elementos de imagem. Sejam os níveis de cinza denotados por r_1, r_2, \dots, r_N e as correspondentes probabilidades de ocorrência P_1, P_2, \dots, P_N . A entropia associada a essa distribuição é dada por:

$$H = \sum_{i=1}^N P_i \log_2 P_i. \quad (6.9)$$

É possível mostrar que, se os elementos de imagem são independentes (o que é uma hipótese pouco realista na prática), o número médio de bits utilizado pelo código de Huffman não pode ser inferior a H .

O procedimento para obter o código de Huffman pode ser descrito através de um exemplo (Pratt, (1978); Rosenfeld and Kak (1982)).

Sejam os elementos de imagem quantizados com 6 níveis, aos quais se associam probabilidades segundo a tabela abaixo:

$$\begin{aligned} P_1 &= p(r_1) = 0,33, & P_4 &= 0,15, \\ P_2 &= p(r_2) = 0,22, & P_5 &= 0,05, \\ P_3 &= p(r_3) = 0,20, & P_6 &= 0,05. \end{aligned} \quad (6.10)$$

O procedimento para obter o código é o seguinte:

Passo 1: Procure na lista de probabilidades os dois níveis com as probabilidades mais baixas. (No caso P_5 e P_6).

Passo 2: Adicione essas duas probabilidades. Forme uma nova lista, com as probabilidades não-alteradas no Passo 1, adicionada da soma gerada nesse passo. Essa nova lista tem um elemento a menos que a anterior.

Passo 3: Volte ao Passo 1 e repita o procedimento até que a lista contenha apenas um elemento.

Esse procedimento gera uma árvore, conforme mostra a Figura 6.2.

A atribuição das palavras de código parte da raiz da árvore binária e atribui 1 a um ramo e 0 ao outro, continuando o procedimento em cada nó de árvore, conforme ilustra a Figura 6.3.

Deve-se observar que a entropia da fonte de sinal pode ser calculada pela Equação 6.9 e vale 2,32 bits/símbolo, enquanto a taxa média de bits é dada por:

$$\begin{aligned} R &= 0,33 \times 2 + 0,22 \times 2 + 0,20 \times 2 + 0,15 \times 3 + 0,05 \times 4 + 0,05 \times 4 = \\ &= 2,35 \text{ bits/elemento} \end{aligned} \tag{6.11}$$

O código de Huffman tem a propriedade de ser unicamente decodificável, no sentido de que uma sequência de bits só pode ser decodificada de uma única maneira. (Além disso ele é um código-prefixo no sentido de que nenhuma palavra do código é um prefixo de qualquer outra). O código de Huffman é ótimo no sentido de que nenhum outro código unicamente decodificável pode ter menor taxa de bits para a mesma fonte de sinais.

Dois problemas podem ser associados ao código de Huffman: 1) é um código de comprimento variável, o que pode exigir algum tipo de armazenamento intermediário entre transmissor e receptor para lidar com a taxa variável de transmissão; 2) o desempenho do código é bastante sensível a variações na estatística da fonte de sinal, o que dificulta sua utilização em imagens que tendem a ser representadas por fontes não-estacionárias.

6.2.2 - Codificação de Elementos Adjacentes

A correlação existente entre elementos adjacentes de imagem pode ser explorada na codificação de pares de pixels. Todavia, se cada elemento de imagem é quantizado em L níveis, é necessário armazenar L^2 palavras de código (ou gerá-las), o que tende a tornar o processo inviável. Se se considerar 3 elementos adjacentes, além da necessidade de armazenamento de L^3 palavras, o aumento de desempenho não é substancial.

Uma alternativa consiste em codificar diferenças entre elementos de imagem adjacentes numa linha ou em linhas adjacentes. Tal codificação tem alguma semelhança com os códigos preditivos a serem discutidos adiante. Se cada elemento tem L níveis de quantização, a diferença entre esse elemento e o adjacente poderá ter $2L-1$ níveis. Todavia, devido à correlação entre esses elementos a probabilidade de ocorrência de grandes valores de diferença é pequena e é possível desenvolver uma variação do código de Huffman, denominado código de Huffman com deslocamento que, ao preço de um pequeno decréscimo de desempenho, pode diminuir consideravelmente o número de palavras de código (Gonzalez and Wintz, 1977).

6.3 - TÉCNICAS IRREVERSÍVEIS

6.3.1 - Modelos Matemáticos

Com o objetivo de alcançar uma taxa maior de compressão de dados é necessário tolerar algum tipo de distorção. Esses métodos usualmente envolvem três estágios (Netravali and Limb, 1980): 1) um estágio inicial no qual é feito algum tipo de transformação da imagem, em geral reversível, com o objetivo de eliminar a redundância estatística; 2) um estágio, usualmente irreversível, em que a acurácia de representação é reduzida, mantendo-se uma qualidade visual da imagem; 3) um estágio reversível no qual a codificação propriamente é realizada, podendo ser utilizados os códigos estatísticos como os de Huffman. A Figura 6.4 ilustra o diagrama de blocos de um processo de codificação de imagens irreversível.

Para o projeto de um tal sistema é usual a caracterização da imagem do ponto de vista estocástico, através de campos aleatórios. Normal

mente é feita a caracterização através dos dois primeiros momentos, envolvendo o valor esperado dos elementos de imagem e sua matriz de covariância. Sob a hipótese de estacionaridade no sentido amplo do campo aleatório, o valor esperado é constante e a covariância entre elementos de imagem são dependentes de sua separação geométrica. Outra hipótese simplificadora é que a função de covariância dos elementos possa ser dada por um produto de Kronecker das matrizes de covariância nas direções horizontal e vertical. A acurácia desse modelo foi investigada por Jain (1976a). Uma outra simplificação matemática frequentemente feita é a utilização de um modelo markoviano para as matrizes de covariância nas direções horizontal e vertical. Deve-se notar, contudo, que um dos problemas básicos na caracterização de imagens reside na não-estacionaridade do campo aleatório, o que pode exigir o uso de algoritmos adaptativos de codificação. O modelo gaussiano para a caracterização da densidade de probabilidade dos elementos de imagem tem sido utilizado, mais talvez por razões de simplicidade analítica do que por realismo.

6.3.2 - Técnicas Preditivas

Os códigos preditivos têm suas raízes na teoria de predição de séries temporais desenvolvida por Wiener na década de 40. A idéia básica é codificar o sinal resultante da diferença entre o elemento de imagem e a predição dele, baseada na mesma linha (predição unidimensional), em linhas adjacentes (predição bidimensional), ou mesmo envolvendo elementos de imagem nos quadros precedentes, no caso de imagens de TV.

Considere-se o caso mais simples de estimação de um elemento de imagem baseado numa operação linear envolvendo o elemento anterior na mesma linha. Suponha-se também por simplificação que a imagem tenha média nula. Sob essa hipótese, pode-se mostrar (Papoulis, 1965) que um preditor linear que minimize o erro médio quadrático, dado por:

$$E = E \{ (f(i,j) - \tilde{f}(i,j))^2 \}, \quad (6.12)$$

terá a forma:

$$\tilde{f}(i,j) = a f(i-1,j). \quad (6.13)$$

Se a média não fosse nula, a Expressão 6.13 deveria ser adicionada uma constante a ser determinada.

Expandindo-se a expressão 6.12 e utilizando a Equação 6.13 obtêm-se:

$$E = E\{f^2(i,j) - 2a E\{f(i,j) f(i-1,j)\} + a^2 E\{f^2(i-1,j)\}}. \quad (6.14)$$

Sob a hipótese de que o campo aleatório que descreve a imagem é estacionário no sentido amplo, a função de correlação $R(.,.,.,.)$ é da forma:

$$R(i,j,k,l) = R(i-k, j-l) \quad (6.15)$$

e daí resulta:

$$E = R(0,0) - 2a R(1,0) + a^2 R(0,0). \quad (6.16)$$

Diferenciando a expressão anterior em relação ao coeficiente a e igualando o resultado a zero, obtêm-se o valor ótimo de a :

$$a = \frac{R(1,0)}{R(0,0)}. \quad (6.17)$$

Este valor é reconhecido como o coeficiente de correlação (ρ) entre elementos de imagem adjacentes, isto é:

$$\hat{f}(i,j) = \rho f(i-1,j). \quad (6.18)$$

Substituindo o valor de $\hat{f}(i,j)$ dado pela Equação 6.18 na Expressão 6.16, obtêm-se o valor mínimo do erro médio quadrático:

$$\begin{aligned} E_{\min} &= R(0,0) \left[1 - \frac{R^2(1,0)}{R^2(0,0)} \right] = \\ &= \sigma^2 (1 - \rho^2), \end{aligned} \quad (6.19)$$

onde σ^2 é a variância da imagem suposta constante.

Quando se faz a predição de um elemento de imagem por uma combinação linear de elementos adjacentes, a determinação dos coeficientes de predição se faz de maneira análoga, envolvendo a solução do chamado sistema de equações normais. Essas equações têm como coeficientes as correlações entre os elementos de imagem envolvidos na predição. Pode-se mostrar (Rosenfeld and Kak, 1982) que, se o campo aleatório que caracteriza a imagem for markoviano, com função de autocorrelação separável, então a predição de um elemento de imagem baseada nos elementos superior, à esquerda e na diagonal superior à esquerda fornece os resíduos de predição ($f(i,j) - \hat{f}(i,j)$) não correlacionados. Habibi (1971) realizou um estudo de preditores de 1ª até 22ª ordem, concluindo que preditores de 3ª ordem, segundo o esquema bidimensional sugerido anteriormente, apresentam o melhor compromisso entre desempenho e complexidade.

O diagrama de blocos de um sistema típico de transmissão e recepção por códigos preditivos está mostrado na Figura 6.5. Os preditores do transmissor e do receptor são iguais. A função desses preditores é atrasar o sinal de entrada por um tempo igual ao intervalo entre amostras e multiplicar pela constante dada pela Equação 6.17.

A alta correlação normalmente existente entre elementos de imagem adjacentes faz com que o histograma do resíduo seja concentrado ao redor do valor zero. A densidade de Laplace tem sido utilizada para modelar esse histograma e um quantizador ótimo pode ser derivado utilizando a técnica de Max, resultando em pequeno espaçamento para pequenos erros de predição e grande espaçamento para grandes erros. A codificação desses níveis pode ser feita então utilizando eventualmente o código de Huffman. Deve-se levar em conta, entretanto, os problemas que resultam de uma taxa variável de transmissão.

Ao sistema de codificação por predição dá-se, de modo geral, a denominação Modulação com Codificação Diferencial por Pulsos (em inglês "Differential Pulse Code Modulation" - DPCM). Um sistema DPCM com um quantizador ótimo de 8 níveis (3 bits) pode produzir imagens com a mesma quali

dade visual que um sistema PCM de 6 a 8 bits, com exceção dos erros em bordas agudas da imagem.

Quando a quantização é feita em 2 níveis (1 bit), o sistema de codificação tem o nome de Modulação Delta (em inglês "Delta Modulation"-DM). O diagrama de blocos de um sistema de Modulação Delta é representado na Figura 6.6. O sinal analógico serve de entrada para um dispositivo que realiza uma diferença. Se o sinal diferença é positivo, a limiarização produz um pulso; caso contrário, é produzido um pulso negativo. Os pulsos positivos e negativos são codificados como dígitos binários e transmitidos. Simultaneamente, os pulsos são utilizados como entrada do integrador na malha de realimentação, que produz uma representação aproximada do sinal do tipo escada. No sistema de recepção os bits são utilizados para reconstruir os pulsos positivos e negativos que são integrados para reconstruir o sinal original. O filtro passa-baixas serve para eliminar componentes espúrias de altas frequências no sinal reconstruído.

A Figura 6.7 ilustra os problemas básicos que ocorrem com o sistema de Modulação Delta: se o passo de limiarização é pequeno, o sistema não é capaz de acompanhar bordas onde a variação do sinal é rápida (sobrecarga). Se, por outro lado, o passo de limiarização é aumentado para compensar o efeito anterior, surge o problema de ruído granular, que resulta de uma oscilação visível entre os dois níveis de quantização em regiões uniformes da imagem. A escolha adequada do passo de limiarização resulta de um compromisso entre esses dois problemas conflitantes.

Esse fato é ilustrado na Figura 6.8c, que é uma situação intermediária entre a Figura 6.8b (sobrecarga) e a Figura 6.8d (ruído granular). Nessas figuras q representa a porcentagem do passo de limiarização em relação ao pico de sinal e e.m.q indica o erro médio quadrático entre a imagem original e a imagem reconstruída.

Problemas análogos de sobrecarga e ruído granular ocorrem na codificação preditiva com maior número de níveis. A Figura 6.9c representa igualmente um compromisso entre as Figuras 6.9b (sobrecarga) e 6.9d (ruído granular) para um DPCM de 8 níveis.

Códigos preditivos bidimensionais oferecem em geral uma pequena redução do 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 visuais de erros de transmissão (a ser discutido adiante) são diminuídos.

Como já foi mencionado anteriormente, imagens tendem a exibir uma forte não-estacionaridade, o que tem motivado o desenvolvimento de códigos adaptativos. Uma possibilidade que tem sido bastante explorada consiste em adaptar os níveis de quantização ao conteúdo da imagem. Se o brilho da imagem 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 granular. Outra possibilidade consiste em associar ao nível de variação tonal da imagem a escolha de um modo de Modulação Delta (pequera variação) ou DPCM (grande variação).

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 diferença entre a predição e o valor real do elemento de imagem, de modo que um erro de transmissão tem a tendência de se propagar.

No caso de preditores unidimensionais ótimos sabe-se (Woods, 1975) que para a maior parte das funções de correlação encontradas em imagens o preditor é estável, o que implica que os efeitos do erro de transmissão tendem a decair. O preditor baseado diretamente no elemento anterior é instável e os erros de transmissão resultam em faixas nas linhas recebidas. Este preditor pode ser estabilizado por uma "integração com fuga", isto é, apenas uma fração do valor anterior é usada para predição. A Figura 6.10 ilustra o efeito de erros de transmissão de um DPCM com 8 níveis e predição pelo elemento de imagem anterior por uma fração igual a 0,9.

No caso de códigos preditivos bidimensionais, a análise é mais difícil. Preditores ótimos bidimensionais para as funções de correlação usuais podem ser instáveis e um critério adequado de estabilidade ainda não é conhecido. Todavia, há preditores bidimensionais estáveis para os quais o efeito visual de distorção é bem menos perceptível que no caso unidimensional. A Figura 6.11 apresenta os efeitos de erros de transmissão para um código preditor bidimensional de 4ª ordem.

A redundância existente entre quadros sucessivos de TV pode ser explorada por técnicas preditivas. Todavia, devido à dificuldade de armazenamento, tais códigos têm sido baseados na predição apenas pelo quadro anterior. Os códigos preditivos entre quadros exibem a deficiência de "sobrecarga temporal" análogo à sobrecarga descrita para os códigos preditivos espaciais. Se uma cena contém objetos de alto contraste que se movem, a reprodução de suas saídas é deficiente, aumentando em passos igual ao intervalo de quantização. Tem havido algum esforço para combinar os códigos preditivos temporais com a técnica de preenchimento de quadro onde apenas os elementos que variam de quadro para quadro são transmitidos.

6.3.3 - Técnicas por Transformação

A codificação por transformação baseia-se na idéia de aplicar uma transformação linear sobre uma imagem ou sobre uma subimagem (para diminuir a carga computacional), de modo a obter componentes transformadas menos correlacionadas do que os elementos de imagem originais. Em seguida, estas componentes são ordenadas segundo sua importância em representar a informação, e as componentes menos representativas são descartadas (ou quantizadas com um número menor de bits) para obter compressão. No receptor a transformação inversa é obtida para recuperar a imagem. A Figura 6.12 ilustra o esquema geral de codificação por transformação.

6.3.3.1 - Transformação Ótima

Em princípio, a transformação ótima deveria tornar os coeficientes independentes, mas estatísticas de ordem superior a 2 são difíceis de obter, de modo que se deve procurar uma transformação que decorrelacione os coeficientes.

A Transformada de Karhunen-Loeve ou de Hotelling, além de possuir esta propriedade, tem também a característica de minimizar o erro médio quadrático de reconstrução quando se descartam os coeficientes. Considerando uma subimagem como um vetor N-dimensional de elementos de imagem, pode-se calcular a matriz de covariância C_x deste vetor de variáveis aleatórias. A Transformação de Karhunen-Loeve é dada pela matriz A cujas linhas são os vetores próprios normalizados C_x . Admitindo que a matriz C_x seja positiva definida, todos os seus valores próprios são positivos e podem ser ordenados. A máxima compactação de energia é feita preservando os coeficientes correspondentes aos vetores próprios associados aos maiores valores próprios, que representam as variâncias dos coeficientes obtidos. O erro médio quadrático de reconstrução é dado pela soma dos valores próprios das componentes descartadas.

O uso da Transformação de Karhunen-Loeve apresenta vários problemas na prática: 1) a não-estacionariedade da imagem implica que a transformação depende do bloco da imagem; 2) a transformação é dependente da caracterização estatística da imagem, o que é frequentemente difícil de obter; 3) a transformada de N elementos de imagem em geral requer N^2 multiplicações, o que pode representar uma carga computacional considerável. Uma exceção é o caso em que a função de correlação da imagem é do tipo markoviano e separável, isto é:

$$E [f(i,j) f(k,l)] = \sigma^2 \rho_1^{|i-k|} \cdot \rho_2^{|j-l|} , \quad (6.20)$$

onde σ^2 é a variância da imagem (supondo-se média nula) e ρ_1 e ρ_2 são respectivamente os coeficientes de correlação na direção vertical e horizontal, respectivamente. Nessas condições, Jain (1976b) mostrou que é possível implementar a Transformação de Karhunen-Loeve por algoritmos de Transformada Rápida de Fourier.

6.3.3.2 - Transformações Sub-Ótimas

Várias transformações subótimas têm sido utilizadas na compressão de imagens. Elas produzem coeficientes menos correlacionados que a imagem original, não dependem da estatística da imagem e podem ser im

plementadas por algoritmos rápidos. Dentre essas transformadas, podem ser mencionadas a Transformada Discreta de Fourier, a de Hadamard, a de Haar, a Transformada Discreta do Cosseno e a Transformada Oblíqua ("slant").

A Transformada Discreta de Fourier pode ser implementada eficientemente através do algoritmo de Transformada Rápida de Fourier (em inglês "Fast Fourier Transform"- FFT). Como no caso das demais transformações, ela é separável, o que implica que a transformada pode ser obtida aplicando sucessivamente transformações unidimensionais sobre as linhas de imagem e, em seguida, sobre as colunas. Assim sendo, a transformada bidimensional pode ser implementada com $2N^2 \log_2 N$ multiplicações e adições (onde N é o número de linhas de colunas da imagem). Deve ser observado que a Transformada de Fourier é assintoticamente equivalente à de Karhunen-Loeve, isto é, os coeficientes são não-correlacionados quando N tende a infinito (Grenander and Szego, 1969).

A Transformada de Hadamard é baseada na matriz de Hadamard (ver Apêndice), que é composta apenas de (+1 e -1). Assim sendo, ela tem a vantagem de poder ser implementada apenas por somas e subtrações. O mesmo acontece com a Transformada de Haar, envolvendo apenas +1, -1 e 0.

A Transformada do Cosseno tem sido objeto de grande interesse recentemente, 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.

A Transformada Oblíqua é baseada numa matriz cujas linhas são formas de onda discretas monotonicamente decrescentes com passo constante, da máxima para a mínima amplitude. Desse modo, ela produz uma representação adequada para variações graduais de brilho numa linha da imagem.

Deve-se observar que todas as transformações mencionadas são unitárias, isto é, $A^{T*} = A^{-1}$ (onde A^{T*} é o complexo conjugado de A^T) e portanto a transformada inversa, que é feita no receptor, é implementada de maneira análoga à no transmissor.

6.3.3.3 - Seleção dos Coeficientes

Depois da transformação, as amostras transformadas são selecionadas para transmissão segundo dois tipos de estratégia usuais:

a) Amostragem por zona, quando apenas os coeficientes que se situam numa região do plano transformado, em geral associados às frequências mais baixas, são codificados para transmissão. Seu desempenho depende do tipo de imagem a ser codificada. A degradação pode ser grande se a imagem contém grandes componentes que são descartadas. No receptor os coeficientes eliminados são usualmente igualados a zero. A Figura 6.13 ilustra uma típica atribuição de bits aos coeficientes de uma codificação por amostragem por zona, em blocos de 16 x 16 elementos de imagem, a uma taxa média de aproximadamente 1,5 bits/elemento.

A utilização da Transformada de Hadamard para a codificação através de amostragem por zona de uma imagem Landsat da cidade de São Paulo está ilustrada na Figura 6.14.

Exemplos da utilização de diversas transformações, utilizando o mesmo tipo de seleção dos coeficientes, estão mostrados na Figura 6.15.

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. Se, por um lado, este método pode melhorar o desempenho da codificação, por outro lado a complexidade pode crescer devido ao fato de o número de coeficientes a ser transmitidos variar com o bloco. Pode-se obter uma taxa de transmissão fixa mediante uma certa perda de desempenho, se forem selecionados os K maiores coeficientes, para um valor fixo de K.

O método de amostragem por limiar pode ser considerado um método adaptativo de codificação, uma vez que os coeficientes a ser transmitidos variam de bloco para bloco. Um outro tipo de adaptação consiste em obter uma medida de atividade espacial do bloco ("índice de atividade") que pode ser medido segundo uma soma ponderada dos valores absolutos dos coeficientes transformados (Gimlett, 1975) e atribuir mais bits para codificar aqueles blocos com um índice de atividade maior.

6.3.3.4 - Quantização e Codificação dos Coeficientes

Após decidir quais os coeficientes que devem ser transmitidos deve-se quantizá-los e codificá-los. A quantização dos coeficientes transformados tem sido baseada em modelos estatísticos, utilizando, por exemplo, o esquema desenvolvido por Max. O primeiro coeficiente, que representa a média dos elementos de um bloco, é geralmente descrito pela densidade de Rayleigh, enquanto os demais são modelados pela densidade gaussiana. A base teórica de tal modelamento é fornecida pelo teorema do limite central, dado ao fato de que os coeficientes das transformadas são o resultado da combinação linear de um grande número de variáveis aleatórias (256 para um bloco de 16 x 16, por exemplo), embora algo dependentes. Os histogramas obtidos experimentalmente tendem a confirmar o modelo teórico.

A atribuição do número de bits a cada coeficiente pode ser feita dividindo o número total de bits pelos coeficientes de modo que o erro médio quadrático seja minimizado. No caso de variáveis gaussianas, a atribuição de bits é feita de modo que o erro médio de quantização de cada coeficiente seja o mesmo (Huang and Schultheiss, 1963; Segall, 1976). Isto implica que os bits sejam atribuídos aos coeficientes proporcionalmente ao logaritmo de suas variâncias.

Com o uso de técnicas de transformação tem sido possível codificar imagens monocromáticas com cerca de 1,0 a 1,5 bits por elemento, com erro médio quadrático menor que 0,5%, enquanto imagens em cores requerem um adicional de 0,5 a 1,0 bits por elemento (Pratt, 1979).

6.3.3.5 - Efeitos de Erros de Transmissão

Os efeitos de erros de transmissão em codificação por transformadas é completamente diferente daquele que ocorre em métodos preditivos. Naquele método a imagem reconstruída é uma combinação linear dos coeficientes transformados, de modo que o efeito de erros em coeficientes distribuem-se sobre todo o bloco. Em geral, os erros nos coeficientes de mais baixa frequência (ou sequência, no caso da Transformada de Hadamard) são mais visíveis do que os erros em coeficientes de frequência mais alta de

vido ao fato de que acima de uma dada frequência a sensibilidade do olho cai com frequência. Além disso, a probabilidade de erro é maior em coeficientes de baixa frequência, porque usualmente eles são codificados com um número maior de bits. O efeito de erros de transmissão é maior na técnica de amostragem por limiar do que na amostragem por zona, devido aos erros que podem ocorrer nos bits que indicam a posição dos coeficientes codificados (Pratt, 1978). A Figura 6.16 ilustra estes efeitos na codificação por transformada oblíqua.

6.3.4 - Codificação Híbrida

Os códigos por transformação e preditivos têm em comum a característica de efetuar uma decorrelação das amostras, quantizando e codificando os coeficientes obtidos. A codificação por transformadas produz um erro médio quadrático de reconstrução menor que os códigos preditivos em baixas taxas de transmissão às custas de uma complexidade maior na implementação.

Como um compromisso entre as vantagens e desvantagens de ambos os sistemas, foi desenvolvido o sistema de codificação híbrida predição - transformação. Têm sido estudados três esquemas básicos (Netravali and Limb, 1980): 1) transformada unidimensional ao longo das linhas da imagem, seguida de DPCM na direção vertical; 2) transformadas bidimensionais de pequena dimensão (4 x 4, por exemplo) e DPCM, sendo usados os coeficientes transformados do bloco precedente na direção horizontal para efetuar a predição; 3) codificação por transformada bidimensional e DPCM temporal entre quadros. O desempenho visual da codificação híbrida intra-quadro é satisfatório em taxas de 0,5 a 1,0 bits/elemento e no caso entre quadros uma boa qualidade pode ser obtida a 0,25 bits/elemento.

6.3.5 - Outros Métodos de Codificação

Este parágrafo descreverá brevemente alguns métodos mais especializados de codificação.

As técnicas de interpolação têm sido usadas também para codificação de imagens pela transmissão de um subconjunto de elementos de imagem, obtendo-se por interpolação os elementos restantes. Vários métodos de interpolação uni e bidimensional têm sido propostos, alguns incluindo esquemas adaptativos.

Pela chamada codificação por contornos, a imagem é separada em duas partes: as bordas e o restante. As bordas podem ser codificadas utilizando algoritmos desenvolvidos para representar imagens binárias como fac-símile de documentos de desenhos (Arps, 1979).

Este último tipo de imagens tem sido codificado também pelos códigos de comprimento de corridas (em inglês "run length coding"). Através do rastreamento da imagem do 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 ela associados.

Finalizando, 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 do número de partes tende a sofrer mudança de quadro para quadro (técnicas de preenchimento de quadro).

REFERÊNCIAS BIBLIOGRÁFICAS

- ARPS, R.B. Binary image compression. In: PRATT, W.K., ed. *Image transmission techniques*. New York, NY, Academic, 1979.
- BERGER, T. *Rate distortion theory: a mathematical basis for data compression*. Englewood Cliffs, NJ, Prentice-Hall, 1971.
- GALLAGER, R.G. *Information theory and reliable communication*. New York, NY, John Wiley, 1968.
- GIMLETT, T.I.. Use of activity classes in adaptive transform image coding. *IEEE Transactions Communications*. COM 23 (7): 785-786, July, 1975.
- GONZALEZ, R.C.; WINTZ, P. *Digital image processing*. Reading, MA, Addison Wesley, 1977.
- GRENDER, V.; SZEGO, G. *Toeplitz forms and their applications*. Berlin e New York, Springer-Verlag, 1969.
- HABIBI, A. Comparison of N^{th} order DPCM encoder with linear transformations and block quantization techniques. *IEEE Transactions Communications Technology*, COM 9 (6): 948-956, Dec. 1971.
- HUANG, J.J.Y.; SHULTHEISS, P.M. Block quantization of correlated gaussian random variables. *IEEE Transactions Communications Systems*, CS-11 (9): 289-276, Sept. 1963.
- HUFFMAN, D.A. A method for the construction of minimum redundancy codes. *Proceedings of the FRE*, 40 (9): 1098-1101, Sept., 1952.
- JAIN, A.K. A fast Karhunen-Loeve transform for a class of random processes. *IEEE Transactions Communications*, COM-24 (9): 1023-1029, 1976b.
- _____. Some new techniques in image processing. In: WILDE, C.O.; BARRET, E., ed. *Image science mathematics*, Wilde, C.O.; Barrett, North Hollywood, Western Periodicals, 1976a. p. 201-223.
- NETRAVALI, A.N.; LIMB, J.O. Picture coding: a review. *Proceedings of the IEEE*, 68(3); 366-406, Mar., 1980.

- PAPOULIS, A. *Probability, random variables and stochastic processes*. New York, NY, Mc Graw Hill, 1965.
- PRATT, W.K. *Digital Image Processing*, New York, NY, John Wiley, 1978.
- _____. (Ed.). *Image transmission techniques*, New York, NY, Academic, 1979.
- ROSENFELD, A.; KAK, A.C. *Digital picture processing*. 2.ed. New York, NY, Academic, 1982.
- SEGALL, A. Bit allocation and encoding for vector sources. *IEEE Transactions Information Theory*, IT-22 (3): 162-169, Mar. 1976.
- SHANNON, C.E. *The Mathematical theory of communication*. Urbana, IL, Univ. Illinois Press, 1949.
- WOODS, J. Stability of DPCM coders for television. *IEEE Transactions Communications*, COM-23 (8): 845-846, Aug., 1975.

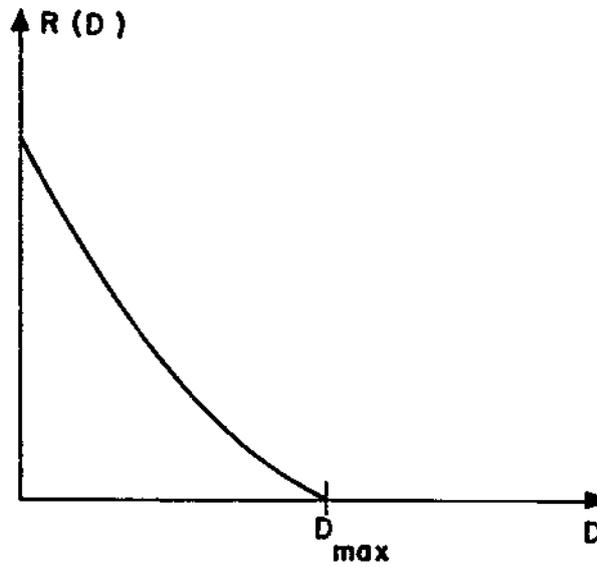


Fig. 6.1 - Uma função de razão de distorção típica.

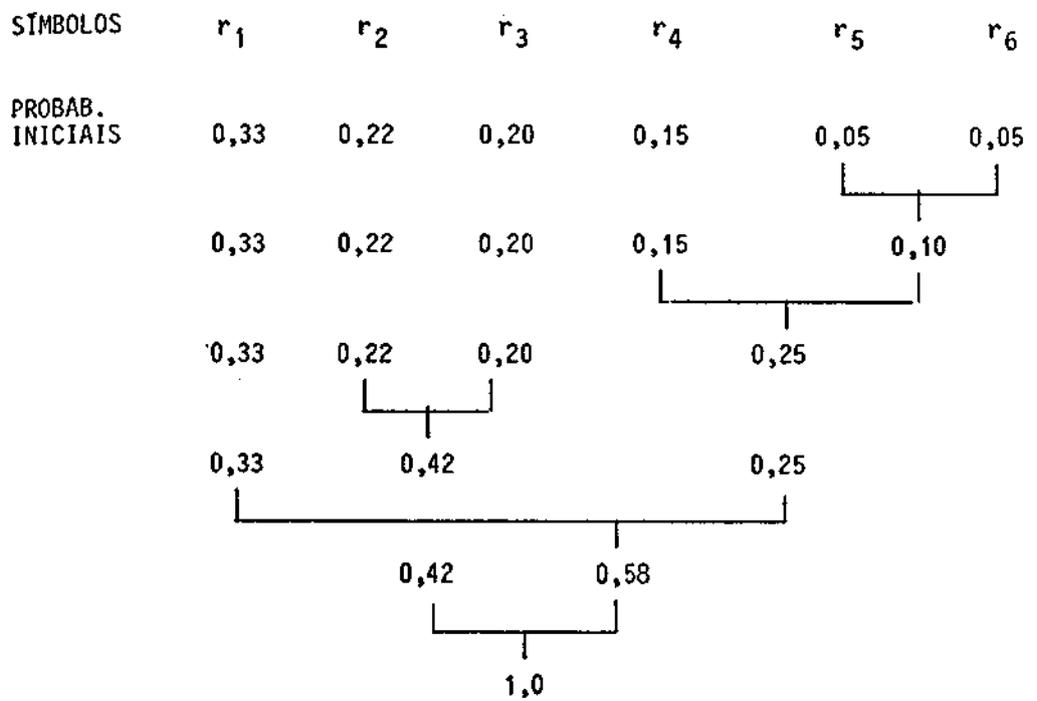


Fig. 6.2 - Árvore para geração do código de Huffman.

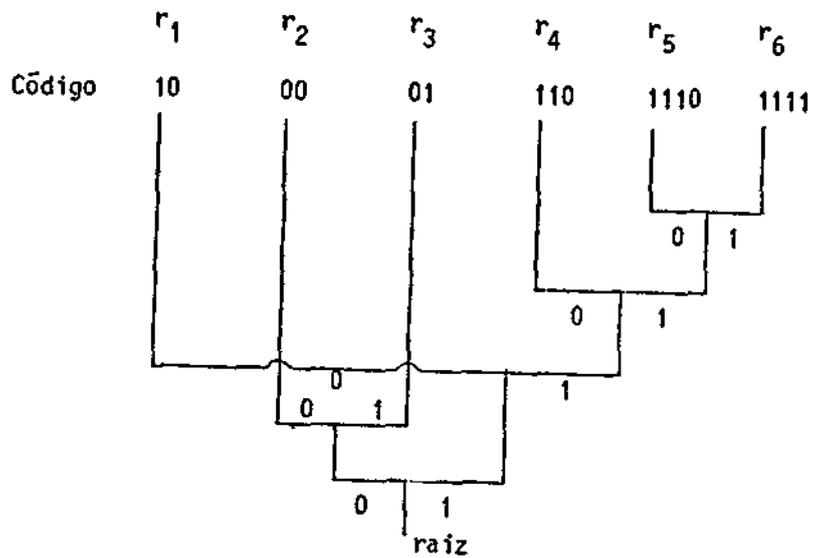


Fig. 6.3 - Utilização da árvore do código de Huffman para determinar as palavras-código associadas aos símbolos da fonte.

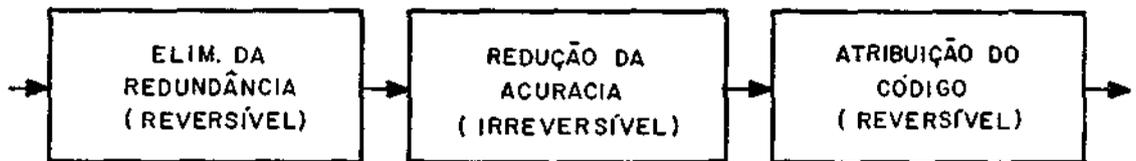


Fig. 6.4 - Diagrama de blocos de um processo de codificação de imagens.

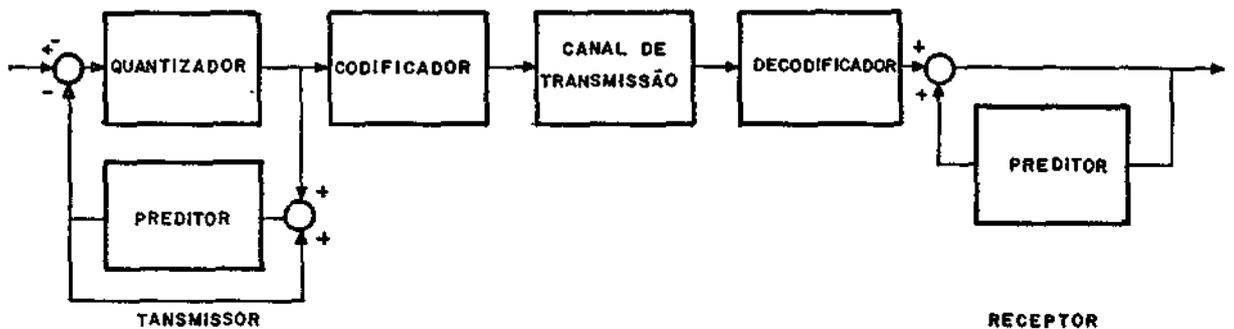


Fig. 6.5 - Diagrama de blocos do sistema de codificação com predição.

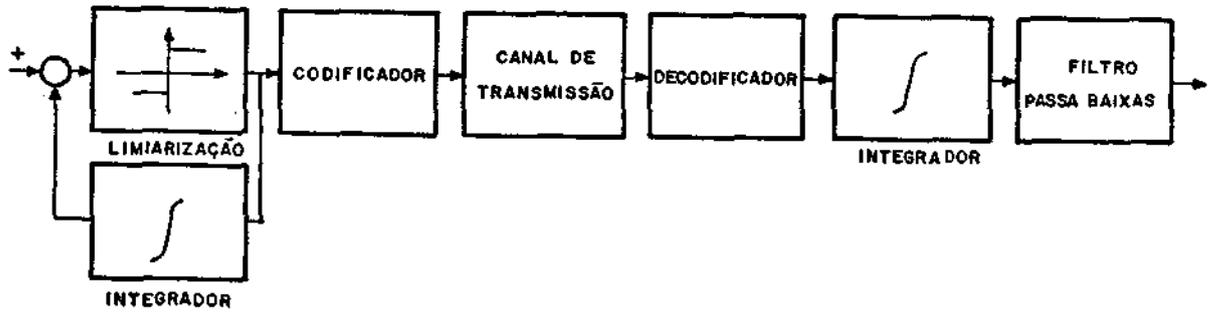


Fig. 6.6 - Diagrama de blocos do sistema de modulação delta.

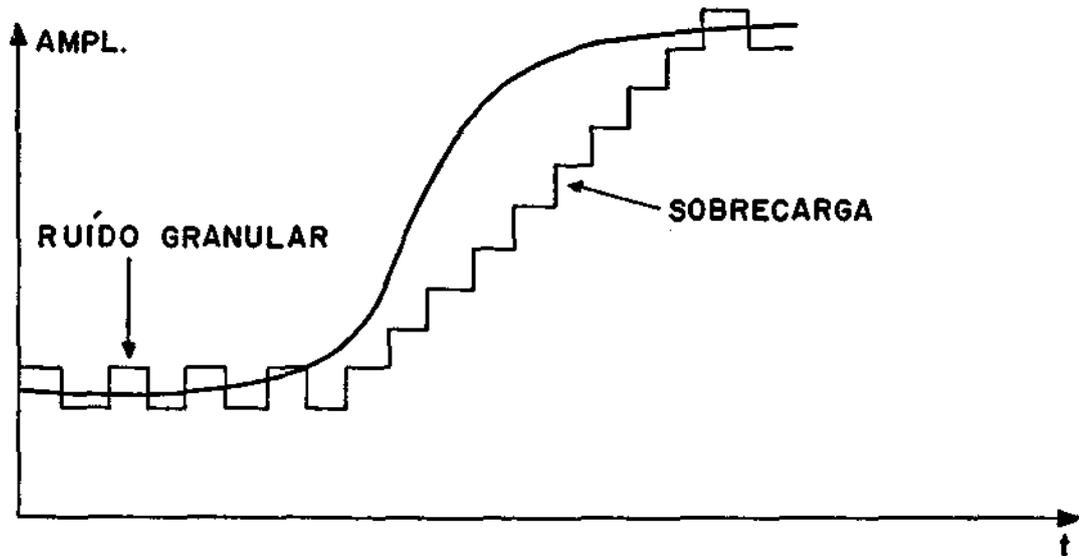


Fig. 6.7 - Distorções no sistema de modulação delta.



(a)



(b)



(c)



(d)

Fig. 6.8 - Exemplos de codificação por DCPM a 3,0 bits/elemento: (a) original; (b) $q = 0,5\%$, e.m.q. = 1,31%; (c) $q = 1,0\%$, e.m.q. = 0,33%; (d) $q = 5,0\%$, e.m.q. = 1,66%.



(a)



(b)



(c)



(d)

Fig. 6.9 - Exemplos de codificação por modulação delta: (a) original; (b) $q = 2,5\%$, e.m.q. = 5.71%; (c) $q = 5,0\%$, e.m.q. = 2.58%; (d) $q = 10,0\%$, e.m.q. = 3.35%.



(a)



(b)

Fig. 6.10 - Efeitos de erro de transmissão para codificação DPCM com 3,0 bits/ elemento: (a) $P(\text{erro}) = 10^{-4}$; (b) $P(\text{erro}) = 10^{-2}$.



(a)



(b)

Fig. 6.11 - Efeitos de erro de transmissão para codificação preditiva por preditor bidimensional de 4ª ordem: (a) $P(\text{erro}) = 10^{-4}$; (b) $P(\text{erro}) = 10^{-2}$.

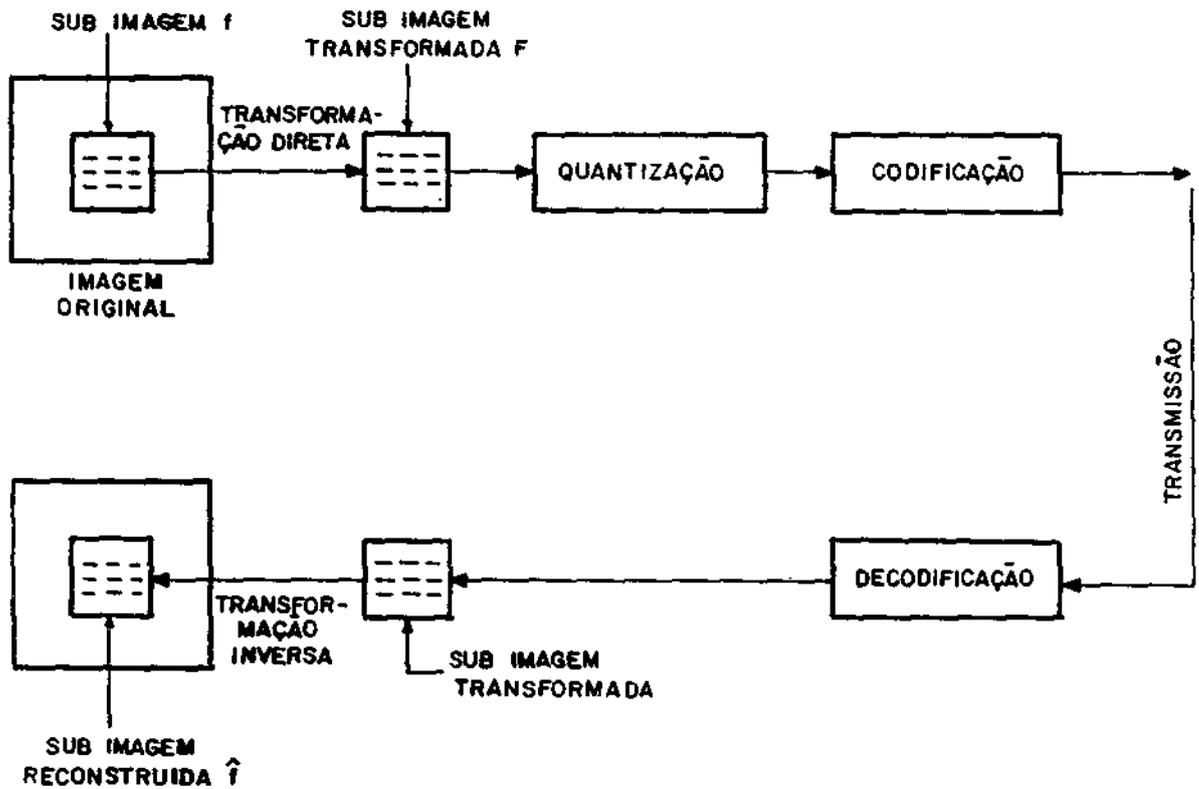
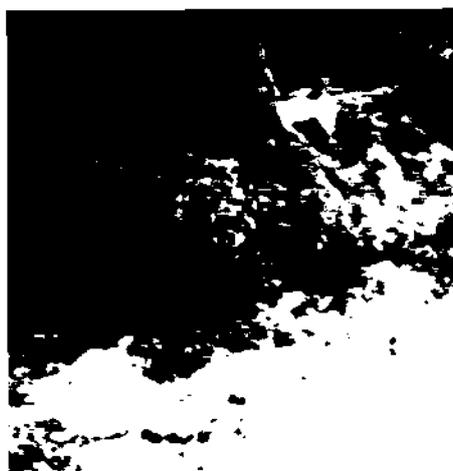


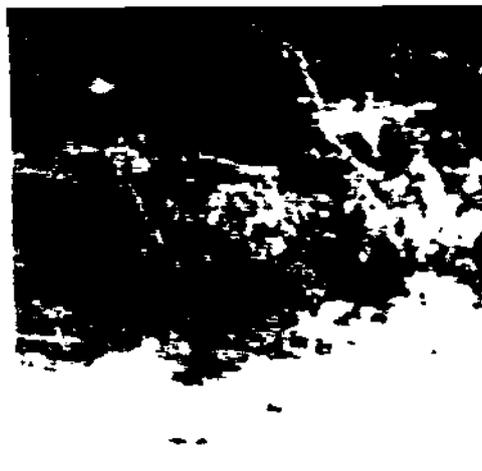
Fig. 6.12 - Esquema geral de um sistema de codificação por transformação.

8	8	6	6	5	5	5	5	4	4	4	4	4	4	4	4	4
8	6	6	4	4	2	2	2	2	2	2	2	2	2	2	2	2
6	6	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2
6	4	4	2	2	2	2	2	2	2	2	2	2	2	2	2	2
5	4	2	2	2	2	2	2	2	2	0	0	0	0	0	0	0
5	2	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0
5	2	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0
5	2	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0
4	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
4	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 6.13 - Atribuição típica de bits em codificação por amostragem zonal, a uma taxa média de aproximadamente 1,5 bits/elemento.



(a)



(b)



(c)



(d)

Fig. 6.14 - Codificação por Transformação de Hadamard - imagem LANDSAT da Região de São Paulo - Banda 4: (a) original; (b) imagem codificada a 1,0 bits/elemento; (c) imagem codificada a 1,5 bits/ elemento (d) imagem codificada a 2,0 bits/elemento.



10-4



10-3



Codificação zonal

10-2

Codificação por limiar

Fig. 6.16 - Efeitos de erro de transmissão com transformada oblãqua em blocos de 16 x 16 elementos, a 1,5 bits/elemento.



(a)



(b)



(c)



(d)



(e)



(f)

Fig. 6.15 - Exemplos de codificação por transformadas, com amostragem zonal em blocos de 16×16 elementos, a 1,5 bits/elemento: (a) original; (b) Hadamard; (c) Haar; (d) oblîqua; (e) cosseno; (f) Karhunen-Loeve.

CAPITULO 7

REGISTRO

Por "registrar" ("casar") duas imagens de uma mesma cena entende-se superpô-las de modo que pontos das imagens que correspondem aos mesmos objetos coincidam espacialmente. A operação de registrar duas imagens é importante por várias razões, entre elas:

- quando se têm imagens obtidas por diferentes sensores e se deseja saber a resposta do mesmo ponto aos diversos sensores;
- quando se têm imagens obtidas em diferentes épocas ou mesmo diferentes instantes de tempo;
- quando se têm imagens tomadas de posições diferentes e se pretende obter informação tridimensional da cena;
- quando se deseja registrar uma imagem obtida por sensores com um mapa.

O primeiro passo a ser dado numa 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 chamados "pontos de controle" nas duas imagens que correspondem ao mesmo objeto. Uma vez que estes pontos são determinados é possível obter funções que mapeiam as coordenadas de uma imagem nas da outra.

O segundo passo consiste na construção da imagem registrada pela aplicação da transformação geométrica. A transformação define as posições dos pontos, e esquemas de interpolação definem o nível de cinza a ser atribuído a cada ponto.

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 dessemelhança entre imagens e computá-la para cada translação possível e tomar a translação de menor medida. É também possível adotar um procedimento semelhante quando há necessidade de rotação: testam-se todas as combinações de translação e rotação até encontrar um par que minimize a medida. Este método, embora possível é, em geral, computacionalmente inviável.

Procedimentos de registro translacional podem ser usados para a determinação dos pontos de controle. Em geral, tomam-se pequenas regiões da imagem que contenham características atípicas e procura-se encontrar estas mesmas características na outra imagem. Este método dá bons resultados se as diferenças (distorções) entre as duas imagens não forem grandes, pelo menos para pequenas regiões.

7.1 - REGISTRO TRANSLACIONAL

Há vários modos de medir a semelhança (ou dessemelhança) entre duas imagens. Basicamente estas medidas computam as diferenças ponto a ponto entre as imagens e combinam estas diferenças num único número.

Sejam I e J duas imagens definidas numa mesma região R; algumas medidas possíveis são:

$$\max_{(x, y) \text{ de } R} [\text{abs}(I(x, y) - J(x, y))], \quad (7.1)$$

$$\sum_{(x, y) \in R} \text{abs}(I(x, y) - J(x, y)), \quad (7.2)$$

$$\sum_{(x, y) \in R} (I(x, y) - J(x, y))^2. \quad (7.3)$$

A Medida 7.1 tem o inconveniente de não levar em conta o número de pontos em que duas imagens diferem. As medidas 7.2 e 7.3 acumulam as diferenças encontradas em cada ponto e a escolha entre elas é mais uma questão de gosto, ou melhor, de conveniência matemática. Embora a Medida 7.3 tenda a penalizar mais as grandes diferenças, esta (soma dos quadrados das dife

renças) apresenta várias propriedades que tornam a computação menos custosa. Nenhuma das três medidas, contudo, modela com fidedignidade a percepção visual humana de diferenças entre imagens.

Caso seja usado $\sum_R (I - J)^2$, como medida adotada, tem-se que:

$$\sum_R (I - J)^2 = \sum_R I^2 + \sum_R J^2 - 2 \cdot \sum_R I \cdot J \quad (7.4)$$

Vê-se, da Equação 7.4, que a semelhança será tanto menor quanto maior for $\sum_R I \cdot J$. Assim, ao invés de $\sum_R (I - J)^2$ como medida de dessemelhança, pode-se tomar $\sum_R I \cdot J$ como medida de semelhança.

O problema tradicional de registro \bar{e} , dada uma imagem I e uma coleção de imagens J_1, J_2, \dots, J_r , achar qual imagem J_s mais se assemelha a I. Caso $\sum_R (J_i)^2 = K$, então tomar $\min_i [\sum_R (I - J_i)^2]$ ou tomar $\max_i [\sum_R I \cdot J_i]$ resulta no mesmo i .

No registro translacional, tem-se uma imagem I com M linhas e N colunas, e uma imagem J com M+A linhas e N+B colunas. Deseja-se saber qual subimagem M x N de J casa melhor com I (conforme a Figura 7.1). A solução é representada por um ponto (x_0, y_0) de J que corresponde ao canto superior esquerdo de I.

Um procedimento para a determinação do registro \bar{e} comparar I com as $(A+1) \cdot (B+1)$ subimagens J_i de J com tamanho M x N. Para cada posição (x', y') calcula-se a medida de semelhança e toma-se o máximo deste valor:

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \cdot J(x+x', y+y') \quad (7.5)$$

Para cada posição (x', y') ter-se-á um valor para o Somatório 7.5. A esta função bidimensional dá-se o nome de "correlação cruzada" de I e J, $C(I, J)$.

Se, para todo (x', y') $\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} J(x+x', y+y')^2$ for constante, o máximo da função $C(I, J)(x', y')$ coincidirá com o mínimo de $(I - J)^2$, que é o que se deseja. Contudo, como isto dificilmente acontece, prefere-se usar a correlação cruzada "normalizada" ao invés da correlação simples:

$$CR(I, J) = C(I, J) / \left(\sum_x \sum_y J(x+x', y+y')^2 \right). \quad (7.6)$$

que apresenta o máximo valor para

$$J(x+x', y+y') = k \cdot I(x, y).$$

O custo computacional da correlação é bastante alto. Este custo pode, todavia, ser menor se a correlação for calculada usando a Transformada de Fourier das imagens, através do resultado:

$$F[C(I, J)] = F^*[I] \cdot F[J],$$

onde $F^*[I]$ denota o conjugado da Transformada de Fourier de I . Para o cálculo da Transformada de Fourier pode ser usado o algoritmo da Transformada Rápida (FFT).

Devido ao alto custo computacional do registro translacional por correlação (mesmo usando Transformada de Fourier) foi proposta uma classe de algoritmos mais eficientes chamados "algoritmos de detecção sequencial de semelhança" (SSDA - Sequential Similarity Detection Algorithms, Barnea and Silverman, 1972).

A idéia básica dos algoritmos SSDA é a seguinte. A cada posição (x', y') de possível registro calcula-se, ponto a ponto, o "erro" (ou diferença) entre I e J' ($J'(x, y) = J(x'+x, y'+y)$). Tão logo o erro acumulado ultrapassa um limiar prefixado, o ponto (x', y') é rejeitado como possível ponto de registro. Os pontos (x', y') para os quais o erro total calcu

lado em toda imagem J' é menor que o limiar são aceitos como possíveis pontos de registro. Um pós-processamento qualquer (por exemplo, escolhe-se o ponto de menor erro) seleciona um único ponto de registro, se desejado. Com isto, através da seleção adequada do limiar, consegue-se eliminar rapidamente a maioria das posições de não casamento. A Figura 7.2 mostra as curvas de crescimento de três posições, duas das quais (A e B) rejeitadas e uma (C) aceita.

Como medida do erro em um dado ponto pode-se tomar a diferença ao quadrado ou o valor absoluto da diferença ou mesmo "normalizar" o valor absoluto da forma:

$$e(x, y) = \text{abs}(I(x, y) - \text{med}(I) - J(x+x', y+y') + \text{med}(J)),$$

onde $\text{med}(I)$ e $\text{med}(J)$ são as médias das imagens I e J' . A cada ponto acumula-se o erro $e(n+1) = e(n) + e(x, y)$, onde $e(0)$ é feito igual a 0.

A escolha do limiar T acima do qual rejeita-se a posição é crítica no desempenho do algoritmo. Se muito baixa pode vir a rejeitar o próprio ponto de registro e, se alta, pode aceitar pontos demais como de registro e consumir mais tempo de execução.

Uma alternativa é tornar o limiar função de n , o número de pontos testados. Deste modo, um limiar $T(n)$, monotonicamente crescente com n (número de pontos), pode eliminar falsos pontos de registro rapidamente e não deixar de aceitar o verdadeiro ponto de casamento, conforme ilustrado na Figura 7.3.

Para o SSDA com limiar crescente é recomendável sortear a ordem em que os pontos são examinados. Dado que pontos vizinhos numa imagem são geralmente correlacionados, se os pontos não forem tomados numa ordem aleatória, é possível que o ponto de registro seja rejeitado devido a diferenças acidentais entre I e J' .

Dois critérios para a escolha de $T(n)$ são:

- 1) A sequência $T(n)$ deve ter a forma aproximada de $e(n)$ (erro acumulado) no ponto de registro, sendo, na realidade, seu limitante superior.
- 2) $T(n)$ deve ter valores iniciais suficientemente grandes para que se possa estabelecer a tendência de crescimento mesmo para pontos longe da região de registro.

A sequência $T(n)$ pode, ainda, ser obtida analiticamente adotando uma distribuição de probabilidade para o erro $e(x, y)$. $e(x, y)$ pode ser visto como a soma de dois erros: um devido à diferença entre as imagens (ruído da imagem) e outro devido ao fato de as imagens estarem fora do registro (ruído de registro). Para o ponto de registro este último erro é nulo. Assim, $T(n)$ deve:

- 1) minimizar a possibilidade de aceitar pontos do registro,
- 2) maximizar a probabilidade de aceitar o ponto de registro,
- 3) minimizar o número esperado de operações.

Como é difícil achar uma sequência que satisfaça os requisitos 1 a 3, podem ser adotadas soluções (subótimas) que, embora são baseadas no ruído da imagem, parecem satisfazer os dois primeiros requisitos. Mais detalhes podem ser encontrados em Barnea e Silverman (1972) ou Mascarenhas e Pereira (1983).

7.1.1 - Registro translacional de imagens binárias

Um caso importante de registro translacional é quando as imagens são binárias. Além do fato de muitas vezes os sensores produzirem imagens binárias, usa-se "binarizar" imagens através de limiares ou da detecção de bordas. Em geral as imagens binárias obtidas apresentam menor conteúdo de ruído da imagem, ou seja, variações de tonalidade, mudanças temporais, etc. Computacionalmente, os métodos de registro de imagens binárias são mais eficientes, isto porque imagens binárias podem ser armazenadas de modo compactado; as operações de multiplicação do cálculo da correlação passam a

ser "e" lógicos que, além de corresponderem a instruções de computador, são realizadas simultaneamente em todos os bits da palavra do computador.

O método mais usado é o da correlação binária. Como visto anteriormente, a correlação é calculada pela fórmula:

$$\sum_x \sum_y I(x, y) \cdot J(x+x', y+y'). \quad (7.7)$$

Como os valores possíveis para $I(x, y)$ e $J(x+x', y+y')$ são 0 e 1, o produto reduz-se ao "e" lógico e a correlação na contagem do número de uns comuns a I e J' . Uma variação do método é contar o número de zeros comuns, o que também pode ser feito de modo eficiente. Uma outra possibilidade é fazer a equivalência lógica entre I e J' , que significa contar os uns e zeros comuns. A Figura 7.4 mostra as superfícies de correlação binária (uns comuns, zeros comuns e unsezeros comuns) entre duas imagens LANDSAT de uma mesma região em torno do ponto de registro separadas de 6 meses. A área de busca (região) $A \times B$ é 50×50 .

7.2 - TRANSFORMAÇÕES GEOMÉTRICAS

Uma transformação geométrica é uma função biunívoca $f: R \times R \rightarrow R \times R$, onde R é o conjunto dos números reais. Costuma-se separar f em suas componentes f_x e $f_y (R \times R \rightarrow R)$ que dão as transformações nos eixos x e y . Dado um ponto (x, y) obtém-se (veja-se a Figura 7.5) (x', y') , onde $x' = f_x(x, y)$ e $y' = f_y(x, y)$. Como a função é biunívoca, existe a função inversa f^{-1} , também separável nas componentes x e y .

O problema pode ser enunciado como: dada uma imagem I (digital) e uma transformação f , obter uma imagem transformada I' . Se a imagem I não fosse digital a solução seria trivial, já que f é biunívoca: o nível de cinza de um ponto (x', y') de I' é igual ao nível de cinza de (x, y) , onde $x = f_x^{-1}(x', y')$ e $y = f_y^{-1}(x', y')$.

Para imagens digitais, não necessariamente o ponto (x, y) , obtido por f^{-1} , é um par de inteiros e, portanto, podem não ter nível de cinza definido, conforme mostra a Figura 7.6.

Uma solução é interpolar o valor do nível de cinza para pontos que caiam fora da grade onde I é definida. Desta forma, na transformação geométrica de imagens, podem-se distinguir dois problemas;

- 1) determinação da função f da transformação,
- 2) escolha do algoritmo de interpolação conveniente.

Estes dois problemas são abordados nas Seções 7.2.1 e 7.2.2, respectivamente.

7.2.1 - Determinação da função de transformação

Uma das finalidades da transformação geométrica é eliminar distorções causadas pelo sistema de imageamento e distorções devidas à posição do sistema. Neste caso, a correção pode ser feita dispondo de um modelo das distorções introduzidas pelo sistema de imageamento e informação precisa da posição do sistema durante a obtenção da imagem. Este tipo de correção é dito "de malha aberta". Os erros remanescentes após a correção são devidos às limitações do modelo e à falta de precisão das informações sobre a posição do sistema de imageamento.

Um exemplo é o caso de imagens obtidas pelo sensor MSS ("multispectral scanner") dos satélites da série LANDSAT de levantamento de recursos terrestres. O satélite coleta imagens da Terra com um conjunto de 6 sensores.

A varredura é feita em direção perpendicular por banda espectral (Figura 7.7).

Entre as distorções introduzidas pelo sistema de imageamento podem-se citar:

- a varredura da cena pelos sensores não apresenta uma velocidade angular constante,

- o atraso de amostragem entre os sensores,
- a superposição entre pontos vizinhos de um mesmo sensor.

Entre as distorções que não são devidas ao sistema de imageamento, as mais importantes são:

- rotação da Terra,
- efeitos de plataforma (desvio de valores de vários parâmetros - altitude, velocidade e atitude - em relação aos valores nominais)
- esfericidade da Terra.

Quando não se dispõe de informação sobre o processo de distorção, a função de transformação deve ser inferida das próprias imagens. O procedimento mais comum é supor que a transformação é (ou pode ser aproximada) por um polinômio. Como a função que é realmente utilizada é a transformação inversa f^{-1} , é mais prático inferi-la diretamente. Deste modo, têm-se (para um polinômio do segundo grau):

$$x = a_0 + a_1.x' + a_2.y' + a_3.x'^2 + a_4.y'^2 + a_5.x'.y',$$

$$y = b_0 + b_1.x' + b_2.y' + b_3.x'^2 + b_4.y'^2 + b_5.x'.y'.$$

A transformação fica perfeitamente conhecida se forem determinados os coeficientes a_i , b_i , $i = 0, 1, \dots$. Esta determinação pode ser feita através de "pontos de controle" existentes nas imagens. Um ponto de controle é um ponto da cena qual se sabe a posição em ambas as imagens. O número de pontos de controle para a determinação dos coeficientes depende do número destes. Para polinômios completos de grau n , o número de coeficientes a determinar é dado por $c(n) = n^2 + 3.n + 2$. Como cada ponto de controle dá origem a duas equações, necessita-se de pelo menos $1/2(n^2 + 3.n + 2)$ pontos de controle.

Uma das transformações mais simples e ainda de utilidade - a transformação plana de semelhança - envolve uma translação, uma rotação e uma escala, conforme mostra a Figura 7.8.

Esta transformação pode ser descrita por polinômios de primeiro grau:

$$x = -a_0 + a_1 \cdot x' + a_2 \cdot y',$$

$$y = -b_0 + a_2 \cdot x' + a_1 \cdot y'.$$

Como há 4 coeficientes a determinar, são necessários dois pontos de controle para especificar esta transformação.

Outra transformação simples - a transformação afim - corresponde a um polinômio do primeiro grau com todos os coeficientes (Equação 7.8). A transformação afim dá um grau maior de liberdade pois permite variar independentemente as escalas nas direções horizontal e vertical:

$$x = a_0 + a_1 \cdot x' + a_2 \cdot y', \quad (7.8a)$$

$$y = b_0 + b_1 \cdot x' + b_2 \cdot y'. \quad (7.8b)$$

Devido a possíveis erros existentes na localização dos pontos de controle e, também, o fato de o polinômio escolhido não ser eventualmente suficiente para modelar a transformação costuma-se tomar mais pontos de controle que o estritamente necessário. Sejam (x_1, y_1) , (x_2, y_2) , ... (x_m, y_m) e (x_1', y_1') , ... (x_m', y_m') as coordenadas dos pontos de controle nas duas imagens I e I' que se deseja registrar. Sejam $a_i, b_i, i=0,1, \dots, n$ os coeficientes do polinômio completo da transformação $(x, y) = f^{-1}(x', y')$. Seja A a matriz $m \times n$ formada por:

$$1 \ x_1' \ y_1' \ x_1'^2 \ x_1' \cdot y_1' \ y_1'^2 \ \dots$$

$$1 \ x_2' \ y_2' \ x_2'^2 \ x_2' \cdot y_2' \ y_2'^2 \ \dots$$

⋮

$$\begin{matrix} \vdots \\ 1 \ xm' \ ym' \ xm'^2 \ ym'^2 \ \dots \end{matrix}$$

Então pode-se escrever:

$$\bar{x} = A \bar{a},$$

$$\bar{y} = A \bar{b},$$

onde $\bar{x}^T \bar{e} [x1, x2, \dots xm]$ $\bar{y}^T \bar{e} [y1, y2, \dots ym]$, $\bar{a}^T \bar{e} [a0, a1, \dots an]$ e $\bar{b}^T \bar{e} [b0, b1, \dots bn]$. Se o número m de pontos for igual a n, e a matriz A tiver inversa, os coeficientes podem ser obtidos por:

$$\bar{a} = A^{-1} \bar{x},$$

$$\bar{b} = A^{-1} \bar{y}.$$

Se o número de pontos de controle m for maior que n, a matriz A não será quadrada e o conjunto de equações é dito sobredeterminado. A solução que minimiza o erro médio quadrático é dada por:

$$\bar{a}^* = A^* \bar{x}, \text{ e}$$

$$\bar{b}^* = A^* \bar{y},$$

onde A* é a "pseudo-inversa" (ou inversa generalizada) dada por:

$$A^* = (A^T A)^{-1} A^T.$$

O erro médio quadrático minimizado por \bar{a}^* e \bar{b}^* é o somatório das distâncias entre os pontos (xi, yi) e (xi^*, yi^*) , $i = 1, \dots m$, onde $\bar{x}^* = A \bar{a}^*$ e $\bar{y}^* = A \bar{b}^*$. Exemplo:

Seja a Tabela 7.1 de pontos de controle. Deseja-se aproximar o mapeamento por uma transformação afim.

TABELA 7.1

PONTOS DE CONTROLE

I		I'	
x	y	x'	y'
0	0	0	1
0	1	0	4
1	1	2	2
1	2	2	3
3	1	3	2

A matriz \bar{e} é dada por:

1	0	1
1	0	4
1	2	2
1	2	3
1	3	2

A pseudo-inversa \bar{e} , portanto,

1.250	-0.250	0.250	-0.250	0
-0.228	-0.162	0.076	0.098	0.217
-0.304	0.284	-0.064	0.132	-0.042

Os vetores de coeficientes são:

REFERENCIAS BIBLIOGRÁFICAS

- ABDOU, I.E.; PRATT, W.K. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of the IEEE*, 67(5): 753-763, 1979.
- AHUJA, N.; HARALICK, R.M.; ROSENFELD, A. Neighbor gray levels as features in pixel classification. *Pattern Recognition*, 12(4): 251-260, 1980.
- ARGYLE, E. Techniques for edge detection. *Proceedings of the IEEE*, 59(2): 285-286, 1971.
- ASKAR, G.P.; MODESTINO, J.W. The contour extraction problem with biomedical applications. *Computer Graphics and Image Processing*, 7(3): 331-355, 1978.
- CHOW, C.K.; KANEKI, T. Automatic boundary detection of the left ventricle from cineangiograms. *Computer Biomedical Research*, 5: 338-410, 1972.
- COLEMAN, G.B.; ANDREWS, H.C. Image segmentation by clustering. *Proceedings of the IEEE*, 67(5): 773-785, 1979.
- COOK, C.M.; ROSENFELD, A. Size detectors. *Proceedings of the IEEE*, 58(12): 1956-1957, 1970.
- DAVIS, L.S. A survey of edge detection techniques. *Computer Graphics and Image Processing*, 4(3): 248-270, 1975.
- DUDANI, S.A.; LUK, A.L. Locating straight-line segments on outdoor scenes. *Pattern Recognition*, 10(3): 145-157, 1978.
- FELDMAN, J.A.; YAKIMOVSKY, Y. Decision theory and artificial intelligence: 1. A semantics-based region analyzer. *Artificial Intelligence*, 5: 349-371, 1974.
- FREI, W.; CHEN, C.C. Fast boundary detection: a generalization and a new algorithm. *IEEE Transactions on Computers*, 26(10): 988-998, 1977.
- GREANIAS, E.C.; MEAGHER, P.F.; NORMAN, R.J.; ESSINGER, P. The recognition of handwritten numerals by contour analysis. *IBM Journal of Research and Development*, 7(1): 14-21, 1983.
- GRIFFITH, A.K. Mathematical models for automatic line detection. *Journal of the ACM*, 20(1): 62-80, 1973.

- MARR, D. Early processing of visual information. *Philosophical Transactions Royal Society of London Series B*, 275: 483-524, 1976.
- MARTELLI, A. Edge detection using heuristic search methods. *Computer Graphics and Image Processing*, 1: 169-182, 1972.
- MARTELLI, A. An application of heuristic search methods to edge and contour detection. *Communications of ACM*, 19(2): 73-83, 1976.
- MILGRAM, D.L. Region extraction using convergent evidence. *Computer Graphics and Image Processing*, 11(1): 1-12, 1979.
- MONTANARI, U. On the optimum detection of curves in noisy pictures. *Communications ACM*, 14(5): 335-433, 1971.
- NAHI, N.E.; JAHANSHAH, M.H. Image boundary estimation. *IEEE Transactions on Computers*, 26(8): 772-781, 1977.
- NEVATIA, R. Locating object boundaries in textured environments. *IEEE Transactions on Computers*, 25(11): 1170-1175, 1976.
- NEVATIA, R. A color edge detector and its use in scene segmentation. *IEEE Transactions Systems, Man and Cybernetics*, 7(7): 820-826, 1977.
- NILSSON, N.J. *Problem-solving methods in Artificial Intelligence*, Mc Graw Hill, New York, 1971.
- OHLANDER, R.; PRICE, K.; REDDY, D.R. Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*, 8(3): 313-333, 1978.
- ROSENFELD, A. Iterative methods in image analysis. *Pattern Recognition*, 10(3): 181-187, 1978.
- ROSENFELD, A.; DAVIS, L.S. Image segmentation and image models. *Proceedings IEEE*, 67(5): 764-772, 1979.
- ROSENFELD, A.; THURSTON, M. Edge and curve detection for visual scene analysis. *IEEE Transactions on Computers*, 20(4): 562-569, 1971.
- SHACHTER, B.J.; DAVIS, L.S.; ROSENFELD, A. Some experiments in image segmentation by clustering of local feature values. *Pattern Recognition*, 11(1): 19-28, 1979.

- TENEBAUM, J.M.; BARROW, H.G. IGS: a paradigm for integrating image segmentation and interpretation. in: CHEN, C., ed. *Pattern Recognition and artificial intelligence*. Academic Press, New York, 1976 p. 472-507.
- THOMPSON, W.B. Textural boundary analysis. *IEEE Transactions on Computers*, 26(3): 272-276, 1977.
- VELASCO, F.R.D. A method for the analysis of gaussian-like clusters. *Pattern Recognition*, 12(6): 381-393, 1980.
- VELASCO, F.R.D.; ROSENFELD, A. Some methods for the analysis of sharply bounded clusters. *IEEE Transactions Systems, Man and Cybernetics*, SMC-10(5): 511-517, 1980.
- WANG, S.; VELASCO, F.R.D.; WU, A.Y.; ROSENFELD, A. Relative effectiveness of selected texture primitive statistics for texture discrimination. *IEEE Transactions Systems, Man and Cybernetics*, 11(3): 360-370, 1981.
- WESZKA, J.G. A survey of threshold selection techniques. *Computer Graphics and Image Processing*, 7(2): 259-265, 1978.
- ZUCKER, S.W. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 5(3): 382-399, 1976.



(a)

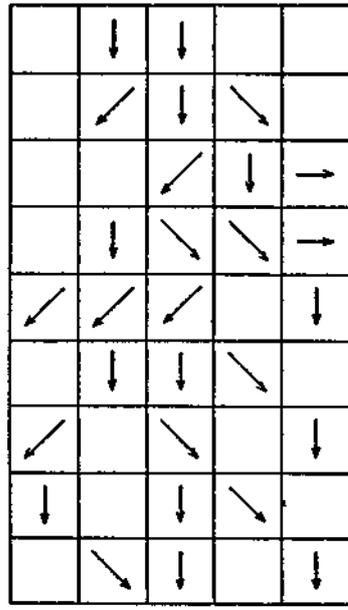


(b)

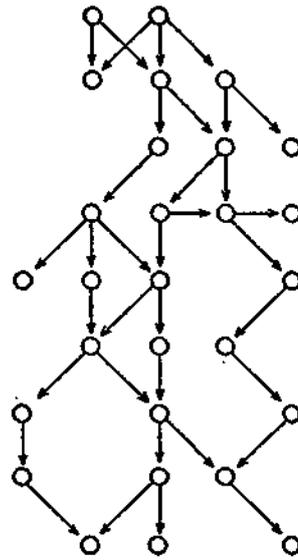
Fig. 8.1 - (a) imagem amplitude de borda; (b) elementos de borda.



Fig. 8.2 - Detecção de bordas segundo a varredura.

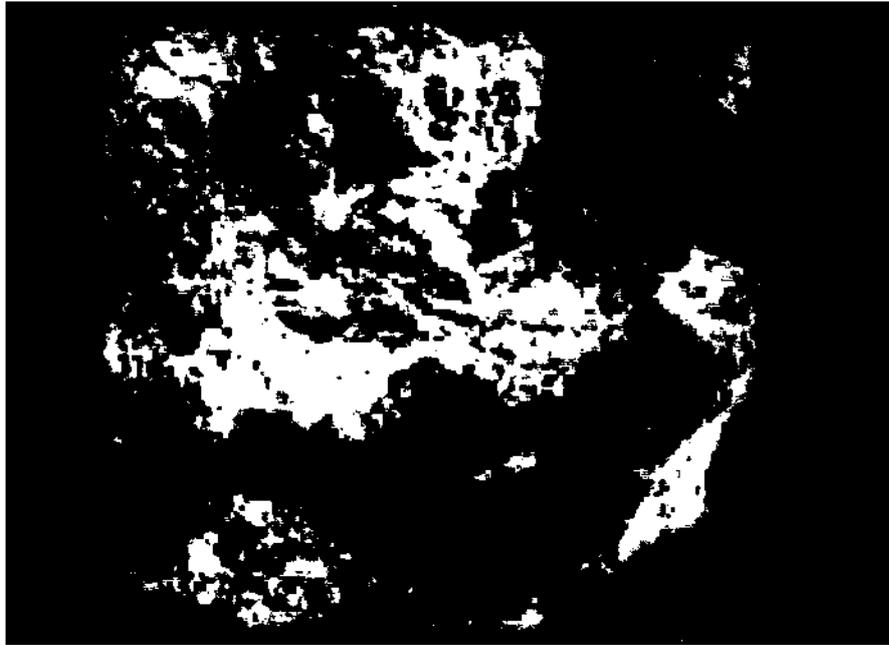


(a)



(b)

Fig. 8.3 - (a) imagem gradiente; (b) grafo correspondente.



(a)

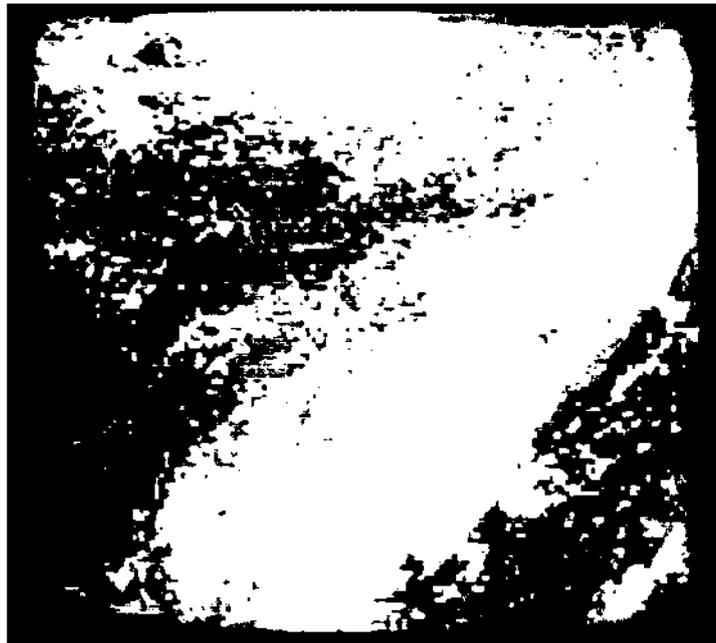


(b)

Fig. 8.4 - (a) imagem; (b) regiões encontradas por crescimento de região se_ gundo a varredura.



(a)



(b)

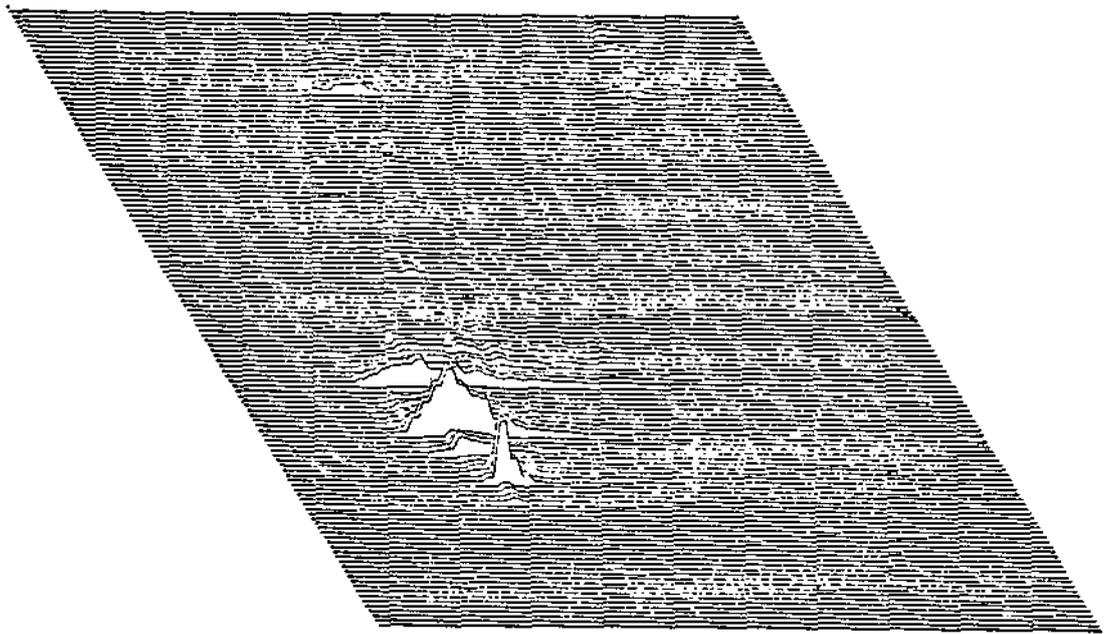


Fig. 8.5 - (a) imagem banda infravermelho; (b) banda visível; (c) histograma bidimensional.



Fig. 8.6 - Imagem classificada.

CAPÍTULO 9

REPRESENTAÇÃO E DESCRIÇÃO

O modo de representar e armazenar uma imagem em computador influi decisivamente no desempenho dos algoritmos que implementam as operações de manipulação e análise. Determina, também, o espaço de memória necessário o que, em várias aplicações, é de suma importância devido ao grande volume de dados que contêm as imagens.

Na descrição, medem-se propriedades dos objetos (regiões ou linhas) da imagem e que caracterizem estes objetos.

Hã dois tipos de imagens que interessa representar: imagens obtidas diretamente de processos físicos e imagens resultantes de processos de segmentação. Estas últimas imagens caracterizam-se pela existência de grandes regiões ligadas com o mesmo valor (classe) para os pontos.

A forma de armazenamento mais comumente adotada é a do armazenamento por linhas, também chamado armazenamento por varredura ("raster"). Em memória principal, este é o método geralmente adotado para o armazenamento de matrizes bidimensionais. Devido ao fato de o acesso aleatório, o tempo de acesso a cada ponto da imagem é o mesmo. Tal não acontece quando a imagem é armazenada em memória secundária de acesso direto. Neste caso, para serem eficientes, os algoritmos devem processar a imagem linha por linha e, se possível, num único passo. Algoritmos como o da Transformada de Fourier tendem a ser ineficientes, não são devido a carga de cálculos, mas também por percorrerem a imagem duas vezes, uma no sentido das linhas e outra no sentido das colunas.

A adequação de um dado esquema de representação pode ser avaliada através de dois fatores principais:

- 1) espaço de armazenamento necessário;
- 2) facilidade de manipulação da imagem sem necessidade de mudança de representação.

Um fator secundário é a facilidade de conversão para uma representação padrão (por exemplo, varredura). A importância deste fator depende da facilidade de operar diretamente com a imagem.

9.1 - ESQUEMAS DE REPRESENTAÇÃO

Como resultado da segmentação de uma imagem digital I tem-se um conjunto de regiões S_1, S_2, \dots, S_n . Uma possível representação para a imagem segmentada é a imagem I_S , onde o valor i de cada ponto indica que o ponto pertence à região S_i . Outra forma de representação é usar n imagens binárias IB_1, IB_2, \dots, IB_n , onde um ponto de IB_i tem valor 1 se o ponto pertence a S_i e tem valor 0 caso contrário. Este modo apresenta a vantagem de poder representar regiões cujas intersecções sejam não-vazias. IB_i é, na realidade, a "função característica" do conjunto S_i . Neste capítulo supõe-se pois, que as imagens que se deseja representar e armazenar são as imagens binárias IB_i que correspondem às regiões.

9.1.1 - Representação por varredura

O esquema de representação por varredura pode ser modificado para economizar espaço utilizando "códigos de comprimento de corrida". Num código de comprimento de corrida, dada uma linha de zeros e uns, indica-se o valor do bit inicial e, em seguida, os comprimentos das subsequências máximas constituídas de um mesmóbit ("corridas"). Caso haja poucas corridas, a economia de memória pode ser substancial; caso haja muitas, pode-se consumir mais memória do que no caso de sequências binárias, pois deve-se armazenar inteiros. Exemplo:

Falta a linha

O código da corrida é 01155121127.

Deve-se notar, entretanto, que em geral, cada linha tem um comprimento diferente, o que torna mais complexo o acesso à imagem.

9.1.2 - Representação de borda

É possível especificar uma região de uma imagem através de sua borda, ou seja, da sequência de pontos que constitui a fronteira da região. A fronteira de uma região S é todo ponto de S que tem um vizinho que não pertence à região. Dependendo se é usada a vizinhança-4 ou a vizinhança-8 as fronteiras são diferentes. Caso seja usada a vizinhança-4 para a região, deve ser usada a vizinhança-8 para a região complementar, ou seja, a composta por zeros.

Seja C um componente ligado de S , ou seja, entre dois pontos quaisquer de C existe um caminho em C que os liga. Pode-se mostrar que se C for adjacente a vários componentes de $-S$ (complemento de S), então só um componente de $-S$ envolve C ; os outros são envolvidos por C . Esta propriedade permite tratar as fronteiras (como será feito) como curvas fechadas. Isto não seria verdade se fossem usadas as mesmas vizinhanças tanto para S quanto para $-S$. Neste capítulo é considerada a vizinhança-4 para S e a vizinhança-8 para $-S$.

Uma sequência de pontos na qual dois pontos consecutivos são vizinhos pode ser representada através da especificação do ponto inicial e de um "código de cadeia". O código de cadeia indica quais os movimentos que devem ser feitos para passar de um ponto ao seu vizinho consecutivo. No caso de vizinhança-8, os possíveis movimentos são:

3	2	1
4	x	0
5	6	7,

onde x é o ponto antes do movimento e $i(0, 1, \dots, 7)$ o ponto após o movimento. O movimento é indicado pelo número i . Exemplo:

A região

A	B	
	C	D
E	F	G
	H	

pode ser representada pela borda ACEHGDBA cujo código de cadeia é (A)7571234, onde A denota a posição inicial da cadeia.

Em lugar do código de cadeia tal como foi apresentado é possível usar o "código de diferença de cadeia", onde os valores representam as mudanças de direção. Assim, a não-mudança é representada por 0; 1 e -1 representam desvios de 45 e -45 graus; e 2 e -2 desvios de 90 e -90 graus. É convencionalmente uma direção inicial, por exemplo, horizontal para a direita. Exemplo:

A região cujo código de cadeia é (A)7571234 no código de diferença de cadeia fica (A)-1-112111.

Para cadeiras com grandes trechos retilíneos uma combinação de código de diferença de cadeia e código de comprimento de corrida resulta em grande economia de memória.

9.1.3 - Representação por árvores quaternárias ("quadrees")

Seja I uma imagem com $N = 2^n$ linhas e colunas. O esquema de árvores quaternárias consiste em dividir a imagem em quatro quadrantes iguais de $2^{(n-1)}$ linhas e colunas. Os quadrantes que estiverem totalmente contidos em S ou -S são rotulados com 1 ou 0, os demais são subdivididos de novo, e assim sucessivamente até resultarem em quadrantes homogêneos. A esta subdivisão da imagem associa-se uma árvore quaternária (cada nó tem 4 ou 0 filhos), onde a raiz corresponde à imagem original, cada nó corresponde a um quadrante, e os filhos de cada nó aos quadrantes em que o nó é dividido, caso não seja homogêneo. Os nós-folha correspondem, portanto, aos quadrantes homogêneos e os demais nós aos quadrantes não-homogêneos. A Figura 9.1 mostra uma

imagem 8x8 (9.1a) e a árvore quaternária associada (9.1b). Os nós-folha brancos correspondem a quadrantes de -S enquanto os pretos a quadrantes de S. A ordem adotada para os quadrantes é dada na Figura 9.2.

O número máximo de níveis de uma árvore quaternária de uma imagem $2^n \times 2^n$ é $n+1$. O espaço de armazenamento é proporcional ao número de nós que, por sua vez, depende da forma e posição da região representada. Regiões compactas tendem a ter menos nós que regiões de forma mais complexa. Mesmo para uma mesma região a árvore depende da posição da região dentro da imagem. Esta é uma propriedade indesejável da representação por árvores quaternárias: pequenas translações podem alterar completamente a aparência da árvore. O esquema de árvores quaternárias não é, portanto, invariante a deslocamento. Compare-o, por exemplo, com o esquema de representação por bordas, onde uma translação implica somente uma alteração do ponto inicial da cadeia.

A representação pode ser estendida para imagens retangulares ou quadrados cujos lados não sejam potências de 2. Para tanto, basta considerar o menor quadrado cujo lado seja uma potência de 2 e inserir o retângulo neste quadrado tomando os pontos fora do retângulo como de -S.

9.1.4 - Conversão entre representações

Nesta seção é visto o problema da conversão entre o esquema de varredura e os demais esquemas (bordas e "quadrees").

a) Bordas a partir de varredura

Para obter a borda de uma região S basta, para cada ponto P de S, fazer a operação booleana $P \& (-P_n \vee -P_s \vee -P_o \vee -P_e)$, onde $\&$ é o "e lógico"; \vee é o "ou lógico"; $-$ é a operação de negação; e P_n, P_s, P_o, P_e são os vizinhos-4 de P. Se o resultado da expressão for 1 então P é de borda. Esta expressão é válida se o conceito de vizinhança para S for vizinhança-4; a extensão para vizinhança-8 é imediata.

A extração da borda de uma região através da operação lógica de um ponto com seus vizinhos só é prática se as operações lógicas puderem ser

feitas em paralelo, para todos os pontos. O Algoritmo 9.1 mostra um procedimento para extrair a borda sequencialmente, a partir de um ponto da borda.

ALGORITMO 9.1 - *Extração sequencial de bordas, a partir da representação de varredura.*

- 1) Faça borda igual a sequência vazia;
- 2) escolha 2 pontos $(P, Q) = (P_0, Q_0)$, tal que P é de S , Q é de $-S$ e P é vizinho-4 de Q ;
- 3) adicione o ponto (P, Q) à borda;
- 4) sejam $R_1 = Q, \dots, R_8$ os vizinhos-8 de P , ordenados no sentido anti-horário;
- 5) seja R_i tal que R_i é de S e não existe $R_j, j < i$, tal que R_j é de S ;
- 6) se $P = P_0$ e $Q = R_k, k < i$, então pare;
- 7) faça $(P, Q) \leftarrow (R_i, R_{i-1})$; adicione (P, Q) à borda;
- 8) vá para 3.

O ponto inicial por onde começa a extração da borda pode ser determinado por uma varredura da imagem: é o primeiro ponto achado(=1) de S , o ponto anterior é certamente de $-S$.

No algoritmo 9.1 foi feita a suposição que S não é formada por um único ponto isolado, pois não se acharia, no passo 5, nenhum vizinho P' de P . Foi suposto, também, que S não toca as bordas da imagem.

A ordenação dos vizinhos de P (passo 4) no sentido anti-horário faz com que as bordas externas sejam percorridas no sentido anti-horário.

rio e as bordas dos buracos no sentido horário. É necessário, também, garantir que todas as bordas (a externa e as internas) tenham sido percorridas. O Algoritmo 9.1 dá o percorrimento de uma única borda.

É necessário comparar cada novo par de pontos com o par de pontos iniciais pois, eventualmente, um ponto S pode se repetir na borda sem que ela tenha sido percorrida totalmente.

Os pares de pontos encontrados ao longo de uma borda são mostrados na Figura 9.3. O "H" indica o ponto Q (de -S) em cada movimento do Algoritmo 9.1.

b) Varredura a partir de Borda

O processo de obter a representação de varredura a partir da borda pode ser feito em dois passos. O primeiro segue a borda marcando com uns os pontos de -S vizinhos aos da borda. Uma vez feito isto, através de um percorrimento da imagem é possível preencher os outros pontos de S com uns e os de -S com zeros.

ALGORITMO 9.2 - *Varredura a partir da borda.*

Passo 1: Rotulação (0 e 1) dos pontos da borda e seus vizinhos de-S.

- 1) Seja P o ponto inicial da borda e P' o ponto seguinte;
- 2) marque P e P' com 1;
- 3) ache Q' vizinho de P anterior a P' no sentido anti-horário;
- 4) $(P, Q) \leftarrow (P, Q')$;
- 5) caso toda a borda e seus vizinhos tenham sido marcados, pare: fim do passo 1;

- 6) ache o novo ponto de borda P' ; marque P' com 1 e todos os pontos anteriores (até Q') com zeros (no sentido horário em torno de P);
- 7) Q' ← ponto anterior a P' ;
- 8) vá para 4.

A execução do Passo 1 é ilustrada pela Figura 9.4 que mostra a sequência de movimentos para rotular a borda e seus vizinhos.

Após executar o Passo 1 para todas as bordas, nenhum ponto não-rotulado é vizinho simultaneamente de 1 e 0. Assim, para preencher o resto de S e $-S$ basta rotular com 1 os pontos não-rotulados vizinhos-4 de uns e com 0 os pontos vizinhos-4 de zeros, o que é feito no Passo 2.

Passo 2: Rotulação dos pontos não rotulados no Passo 1.

- 1) Percorra a imagem segundo a varredura e para cada ponto P , se P for vizinho-4 de algum ponto rotulado, faça o rótulo de P igual ao deste vizinho.

c) Árvores quaternárias a partir de varredura

Na definição do esquema de árvores quaternárias está implícito um método para a obtenção da árvore a partir de representação por varredura. Este método consiste na subdivisão da imagem em quadrantes e no exame da homogeneidade destes quadrantes; os que não forem homogêneos são subdivididos, e assim por diante. Contudo este método tem o inconveniente de o acesso à imagem não ser por linha.

É possível construir a árvore quaternária sequencialmente, linha por linha. No algoritmo proposto por Samet (1981b) constrói-se a árvore passo a passo, partindo da árvore construída por um único nó (branco). A qualquer instante de tempo existe uma árvore válida, onde supõe-se que os pontos não-processados são brancos (de $-S$). A cada ponto da imagem que pertence a S (ponto preto) adicionam-se novos nós à árvore e, se ambas as coordenadas deste ponto forem pares, testa-se se podem ocorrer fusões (quatro

nós pretos filhos de um mesmo nó pai). Ao ocorrer uma fusão, outras podem ser possíveis; o algoritmo deve fazê-las todas antes de passar ao próximo ponto.

d) Varredura a partir de árvores quaternárias

Dada uma árvore quaternária pode-se construir a representação de varredura simplesmente percorrendo a árvore e criando blocos de uns e zeros de acordo com os nós-folha visitados da árvore. Para produzir a imagem, linha por linha (e não bloco por bloco), é necessário um procedimento bem mais complexo que é mostrado no Algoritmo 9.3. O número de vezes que cada nó é visitado no Algoritmo 9.3 é igual ao número de linhas contidas no bloco representado pelo nó.

ALGORITMO 9.3 - *Varredura a partir de árvores quaternárias.*

Para cada linha i da imagem faça:

- 1) $j \leftarrow 1$;
- 2) visite o nó folha A que contém o ponto (i, j) ;
- 3) ache o tamanho ($=2^k$) do bloco representado por A ;
- 4) faça a linha $[j : j + 2^k - 1] \leftarrow$ rótulo de A ;
- 5) $j \leftarrow j + 2^k$; se $j < N$ (N é o comprimento da linha), então vá para 2.

9.2 - OPERAÇÕES COM IMAGENS

Nesta seção são estudadas algumas operações feitas com imagens temáticas. De um modo geral, os algoritmos que implementam estas operações são mais simples quando o esquema de representação usado é o de varredura. Em alguns casos é dado também os algoritmos em outras representações.

9.2.1 - Encolhimento e expansão

As operações de encolhimento e expansão, também chamadas "erosão" e "dilatação", são bastante usadas para pós-processar a imagem obtida da segmentação (supressão de ruído e suavização de contorno, por exemplo).

Encolher uma região S consiste em retirar de S sua borda; a expansão consiste na adição à borda de $-S$. A operação de encolhimento, se repetida várias vezes, leva ao desaparecimento de S . Seja $S(-k)$ a região obtida encolhendo S k vezes, seja $S(k)$ o resultado de k expansões de S , e $S(0)$ por definição o mesmo que S . Deve-se observar que as operações de encolhimento e expansão não são comutativas, ou seja, $S(m)(-n)$ não é necessariamente igual a $S(-n)(m)$ nem tampouco igual a $S(m-n)$. Contudo dois resultados válidos são:

$$S(m-n) \subseteq S(m)(-n) \text{ e}$$

$$S(-n)(m) \subseteq S(m-n),$$

onde $A \subseteq B$ se e somente se todo ponto de A pertence também a B . Em particular para $(m=n)$ tem-se que $S(-m)(m) \subseteq S(m)(-m) \subseteq S$.

As operações de encolhimento e expansão podem ser generalizadas considerando vizinhanças arbitrárias $V(P)$ de um ponto P . Neste caso, a expansão de uma região S é o conjunto $\bigcup_{P \in S} V(P)$ e o encolhimento o conjunto

$$\{P \in S: V(P) \cap -S = \emptyset\}.$$

A Figura 9.5 (b, c, d) mostra o resultado de várias contrações e expansões da região da Figura 9.5a. Note-se o desaparecimento do "ruído" em 9.5b e a suavização conseguida em 9.5c. Além disto, a região que estava ligada em 9.5a apresenta-se não ligada em 9.5d.

9.2.2 - Operações lógicas em regiões

Se duas regiões S e T forem representadas por imagens binárias, a obtenção de $-S$, $S \cup T$ e $S \cap T$ é relativamente simples: basta fazer

as operações lógicas de negação, conjunção e disjunção bit a bit em todos os pontos das imagens binárias. Em outros esquemas que não o de varredura, estas operações não são triviais. Particularmente, se o esquema de representação através de bordas for usado, os algoritmos para a obtenção destes conjuntos podem ser extremamente complexos.

Para árvores quaternárias a construção da árvore de $\neg S$ a partir da de S pode ser feita trocando, em S , os nós-folha pretos (comuns) por brancos e vice-versa.

A construção de $S \cup T$ é feita percorrendo, simultaneamente, as árvores de S e de T . As subárvores de S e T iguais são copiadas em $S \cup T$; se S ou T tem um nó preto onde o outro tem cinza (nó não-folha), $S \cup T$ recebe um nó preto.

ALGORITMO 9.4 - *União de duas regiões representadas por árvores quaternárias:*

União (raiz de S , raiz de T , raiz de $S \cup T$) o procedimento união (x, y, z), onde x, y e z são nós, é dado por:

- 1) Se x ou y é preto, então $z \leftarrow$ preto;
- 2) senão, se x é branco, então $z \leftarrow y$;
- 3) senão, se y é branco, então $z \leftarrow x$;
- 4) senão, faça:
 - 4.1) $z \leftarrow$ cinza;
 - 4.2) união (x_i, y_i, z_i), $i=0, 1, 2, 3$, onde x_i, y_i, z_i são os filhos de x, y e z , respectivamente.

O algoritmo para a obtenção de $S \cap T$ é semelhante ao Algoritmo 9.5. Ao percorrer dois nós, se um deles for branco, o nó correspondente de

$S \cap T$ também o será; se um deles for cinza e o outro preto, então o não será igual ao cinza; finalmente, se ambos forem cinzas, deve-se percorrer os filtros de ambos os nós.

Uma operação frequentemente realizada é determinar, dado um ponto (x, y) , se ele pertence ou não à região S . Esta operação é realizada, para árvores quaternárias, pelo Algoritmo 9.5. Neste algoritmo os quadrantes N, NE, SO e SE são numerados 0, 1, 2 e 3 (veja-se a Figura 9.1); usam-se também, para x e y , as suas representações binárias.

ALGORITMO 9.5 - *Pertinência de ponto à região.*

- 1) $i \leftarrow 1$;
- 2) sejam b_x e b_y os i -ésimos bits da representação binária de x e y (da esquerda para a direita);
- 3) visite o quadrante Q cujo número é $b_x b_y$ (binário);
- 4) se Q é preto, pare: o ponto pertence à região;
- 5) se Q é branco, pare: o ponto não pertence à região;
- 6) se Q é cinza, então faça $i \leftarrow i+1$; vá para 2.

Supõe-se, no Algoritmo 9.5, que S não é nem a região vazia nem a região formada por todos os pontos da imagem. Para tratar estes casos, é suficiente incluir um teste inicial: se a raiz for branca o ponto não pertence à região, se for preta, então o ponto pertence.

9.3 - PROPRIEDADES GEOMÉTRICAS

Após a segmentação da imagem o passo seguinte é, em geral, extrair das regiões propriedades que as descrevem tão completamente quanto possível. Em níveis mais altos o processamento é feito usando as descrições das regiões e não elas mesmas.

Hã dois tipos de propriedades que caracterizam uma regiãõ: as propriedades "radiomêtricas" que têm a ver com o nível de cinza dos pontos que compõem a regiãõ, e as propriedades "geomêtricas" que se referem à forma e posição da regiãõ. Entre as propriedades radiomêtricas estão as diversas estatísticas de níveis de cinza (média, desvio padrão, máximo, mínimo, etc.). Nesta seção são estudadas sã as propriedades geomêtricas das regiões consideradas individualmente. Não são vistas, portanto, as propriedades geomêtricas de várias regiões quando tomadas em conjunto, ou seja, as relações espaciais. Supõe-se também que as regiões são compostas de um único componente ligado; as extensões para vários componentes são, na maioria das vezes, simples.

9.3.1 - Tamanho

9.3.1.1 - Área

A área de uma regiãõ é definida como o seu número de pontos. O cálculo da área pode ser feito através de algoritmos simples na maioria as representações. No caso da representação por varredura um único percurso da imagem é suficiente para contar os pontos que pertencem e os que não pertencem à regiãõ.

Se a regiãõ S , sem buracos, for representada pela sua borda, pode-se achar a área integrando a regiãõ envolvida pela borda. Seja $C = (P) i_1 i_2 \dots i_n$ o código de borda de S , onde P é o conjunto inicial e i_j pertence a $[0, 7]$. Então a área de S pode ser calculada pelo Algoritmo 9.6.

ALGORITMO 9.6 - *Cálculo da área a partir da representação por bordas.*

- 1).Área \leftarrow 0;
- 2) leia ponto inicial do código (x_0, y_0); $y \leftarrow y_0$;
- 3) leia código (i); $i_1 \leftarrow i$; $i_0 \leftarrow i$;
- 4) leia código (i); $i_2 \leftarrow i$;

- 5) execute a ação da Tabela 9.1 para a entrada (i1, i2);
- 6) $i1 \leftarrow i2$;
- 7) se o código não acabou, $v \leftarrow 4$;
- 8) execute a ação da Tabela 9.1 para (i1, i0);

TABELA 9.1

AÇÕES A SEREM REALIZADAS A CADA PASSO DO ALGORITMO 9.6

y_1 x_2	0/7	1/2	3/4	5/6
0/1	a	a;b	a;b;c	a;b;c;d
2/3	b;c;d;a	b	b;c	b;c;d
4/5	c;d;a	c;d;a;b	b	c;d
6/7	d;a	d;a;b	d;a;b;c	d

Ações:

a: $\bar{a}rea \leftarrow \bar{a}rea - y$.

b: $y \leftarrow y + 1$.

c: $\bar{a}rea \leftarrow \bar{a}rea + y$.

d: $y \leftarrow y - 1$.

A Tabela 9.2 mostra os valores de y e da $\bar{a}rea$ nos vários passos do Algoritmo 9.6, para a região:

A B
C D
E F G
H,

cujo código de borda é (A)7571234, supondo que as coordenadas de A são (5, 5).

TABELA 9.2

PASSOS DO ALGORITMO 9.6

CÓDIGO	i1	i2	AÇÃO	ÁREA	y
-	-	-		0	5
7	7	-		0	5
5	7	5	d	0	4
7	5	7	c;d;a	1	3
1	7	1	d;a;b	-1	3
2	1	2	a;b	-4	4
3	2	3	b;c	1	5
4	3	4	b;c	7	6
-	4	7	c;d;a	8	5

Na representação por árvores quaternárias o cálculo da área é feito percorrendo a árvore e, a cada nó preto, somando $2(h-n)$ à área, onde h é a altura total da árvore e n é o nível do nó (por convenção, o nível da raiz é 0, de seus filhos é 1, etc.).

9.3.1.2 - Perímetro e comprimento de arco

O perímetro de uma região é o comprimento total de sua borda. O comprimento da borda, ou em geral de uma curva digital qualquer é obtido somando 1 a cada movimento horizontal ou vertical e $\sqrt{2}$ aos movimentos na diagonal.

Se a região estiver representada por varredura ou por árvore quaternária é necessário percorrer a imagem ou a árvore para extração da borda. Para árvores quaternárias, que é o caso menos simples, deve-se visitar cada nó preto e verificar se seus vizinhos têm ou não a mesma cor e qual

a contribuição do nó para o perímetro. Mais detalhes podem ser encontrados em Samet (1981c).

9.3.1.3 - Extensão, seção cruzada e assinatura

A altura de uma região pode ser definida como a distância entre o ponto mais baixo e o mais alto da região da imagem. Analogamente, a distância entre o ponto de S mais à esquerda e o ponto mais à direita é a "largura" de S. Estas definições podem ser generalizadas com a noção de "extensão" de uma região numa dada direção que vem a ser a distância entre os pontos extremos de S, medida paralelamente à direção. A maior extensão de S em qualquer direção dá-se o nome de "diâmetro".

A "seção cruzada" de S segundo uma direção consiste na interseção de S com a reta nesta direção. Cada seção cruzada é, pois, uma linha composta de zeros e uns.

Outra propriedade relacionada à extensão e à seção cruzada é a "assinatura" de S segundo um dado eixo. Define-se assinatura de S segundo o eixo x como a função $a(x) = \sum_y I(x, y)$, onde I é a imagem. A assinatura segundo o eixo y é definida analogamente.

9.3.2 - Forma

9.3.2.1 - Complexidade

A complexidade de uma região pode ser definida, informalmente, como o número de palavras necessárias para descrever a região completamente. Devido aos inconvenientes operacionais desta definição, preferem-se medidas indiretas da complexidade. Uma das medidas mais populares para complexidade é a razão p^2/A (perímetro ao quadrado dividido pela área). Esta medida é minimizada para uma região "circular". A Figura 9.6 mostra algumas regiões e suas medidas de complexidade. A medida p^2/A não leva em conta fatores de complexidade, tais como simetria, periodicidade, etc.

9.3.2.2 - Alongamento

O "alongamento" de S pode ser definido como a razão entre o diâmetro (i.e., sua maior extensão) e a extensão na direção perpendicular ao diâmetro. Definido deste modo, o alongamento é sempre maior que 1.

Existem outras definições alternativas para alongamento. Uma delas considera uma aproximação da região para uma forma padrão como retângulo ou elipse e define alongamento como a razão entre o lado maior e o menor (se for retângulo) ou a razão entre o eixo maior e o eixo menor (se elipse).

Ainda outra definição considera alongamento como a razão A/t^2 , onde A é a área da região e t é o dobro do número de passos de encolhimento necessários para fazer a região desaparecer. Esta medida, embora melhor que as anteriores para regiões alongadas e curvas, é pouco confiável para pequenos valores de t. A Figura 9.7 mostra valores de alongamento para duas medidas e regiões. Se for usada m_1 (razão entre extensões) a região mostrada na Figura 9.7a é mais alongada que a da Figura 9.7b; se usada $m_2(A/t^2)$ o resultado é o contrário.

9.3.2.3 - Convexidade

Uma região definida no plano real é "convexa" se, dados dois pontos quaisquer da região, o segmento de reta que os une está completamente contido na região. Para imagens digitais, embora uma definição análoga possa ser usada, deve-se ter cuidado com os efeitos de quantização. Em particular, a linha reta digital entre dois pontos não é definível de modo único.

Um código de cadeia representa um segmento de "linha reta digital" se:

- 1) no máximo duas inclinações diferentes aparecem na cadeia;
- 2) pelo menos um dos códigos ocorre em corridas de comprimento 1;

- 3) o outro código ocorre em corridas de, no máximo, dois comprimentos (ã excessão, possivelmente, das extremidades do segmento).
- 4) se hã dois comprimentos, estes diferem de, no máximo, 1.

Pode-se mostrar que, entre dois pontos quaisquer, hã pelo me nos um segmento de reta digital. É verdade também que pode haver mais que um.

Uma definição de convexidade talvez mais fácil de ser usada é a que considera uma região como convexa se a "curvatura" da borda nunca mu da de sinal. Curvatura é comumente definida como a taxa de variação da "in clinação" que pode ser definida, num ponto P, como a média dos ângulos dos elementos de código de cadeia numa vizinhança de k códigos centrada em P.

REFERÊNCIAS BIBLIOGRÁFICAS

- AHUJA, N.; DAVIS, L.S.; MILGRAM, D.; ROSENFELD, A. Piecewise approximation of pictures using maximal neighborhoods. *IEEE Transactions on Computers*, 27(3):375-379, 1978.
- ALEXANDRIDIS, N.; KLINGER, A. Picture decomposition, tree data structures, and identifying directional symmetries as node combinations. *Computer Graphics and Image Processing*, 8(1):43-77, 1978.
- ARCELLI, C.; LEVIALDI, S. Concavity extraction by parallel processing. *IEEE Transactions Systems, Man and Cybernetics*, 1:394-396, 1971.
- BLUM, H.; NAGEL, R. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10:167-180, 1978.
- DAVIS, L.S. Understanding shape: angles and sides. *IEEE Transactions Computers*, 26(3):236-242, 1977.
- Understanding shape: symmetry. *IEEE Transactions Systems, Man and Cybernetics*, 7(3):204-21, 1977.
- DYER, C.R. Computing the Euler number of an image from the its quadtree. *Computer. Graphics and Image Processing*, 13(3):270-276, 1980.
- DYER, C.R.; ROSENFELD, A. Thining algorithms for grayscale pictures. *IEEE Transaction Pattern Anal. Machine Intelligence*, 1(1):88-89, 1979.
- FREEMAN, H. Computer processing of line-drawing images. *Computing Surveys*, 6:57-97, 1974.
- FREEMAN, H.; SHAPIRO, R. Determining the minimum-area encasing rectangle for an arbitrary closed curve. *Communications ACM*, 18:409-413, 1975.
- HORALICK, R.M. A measure for circularity of digital figures. *IEEE Transactions Systems, Man and Cybernetics*, 4(4):394-396, 1974.
- HUNTER, G.M.; STEIGLITZ, K. Operations on images using quad-trees. *IEEE Transactions Pattern Analysis Intelligence*, 1(2):145-153, 1979.
- MONTANARI, U. Continuous skeletons from digitized images. *Journal ACM*, 16:534-549, 1969.
- NEVATIA, R.; BINFORD, T.O. Description and recognition of curved objects. *Artificial Intelligence*, 8(1):77-98, 1977.

- NORDBECK, S.; RYSTEDT, B. Computer cartography- point-in-polygon programs. *BIT*, 7:30-64, 1967.
- PAVLIDIS, T. A review of algorithms for shape analysis. *Computer Graphics and Image Processing*, 7(2):243-258, 1978.
- PFALTZ, J.L.; ROSENFELD, A. Computer representation of planar regions by their skeletons. *Communications of the ACM*, 10:119-122, 1967.
- ROSENFELD, A. Digital straight line segments. *IEEE Transactions on Computers*, 23(12):1264-1269, 1974.
- ROSENFELD, A. *Picture language: formal models for picture recognition*. New York, Academic, 1979.
- ROSENFELD, A.; PFALTZ, J.L. Distance on digital pictures. *Pattern Recognition*, 1:33-61, 1968.
- SAMET, H. Connected component labelling using quadtrees. *Journal ACM*, 28:487-501, 1981a.
- An algorithm for converting rasters to quadtrees. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 3(1):93-95, 1981b.
- Computing perimeters of images represented by quadtrees. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 3:683-687, 1981c.
- YOUNG, I.T.; WALKER, J.E.; BOWIE, J.E. An analysis technique for biological shape. *Information and Control*, 25:357-370, 1974.
- ZUSNE, I. *Visual perception of form*. New York, 1970.

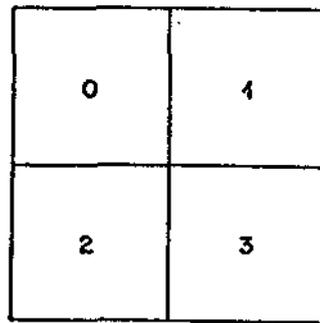
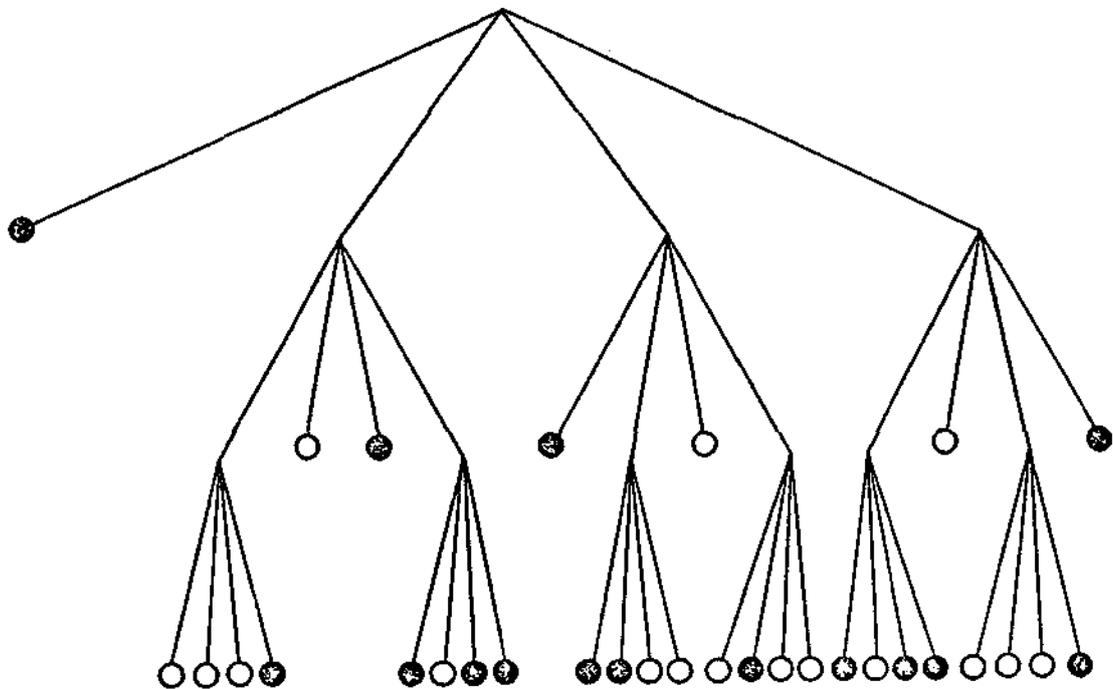


Fig. 9.1 - Ordem dos quadrantes.

1 1 1 1 0 0 0 0
1 1 1 1 0 1 0 0
1 1 1 1 1 1 1 0
1 1 1 1 1 1 1 1
1 1 1 1 1 0 0 0
1 1 0 0 1 1 0 0
0 0 0 1 0 0 1 1
0 0 0 0 0 0 1 1

(a)



(b)

Fig. 9.2 - Imagem binária (a) e árvore quaternária (b).

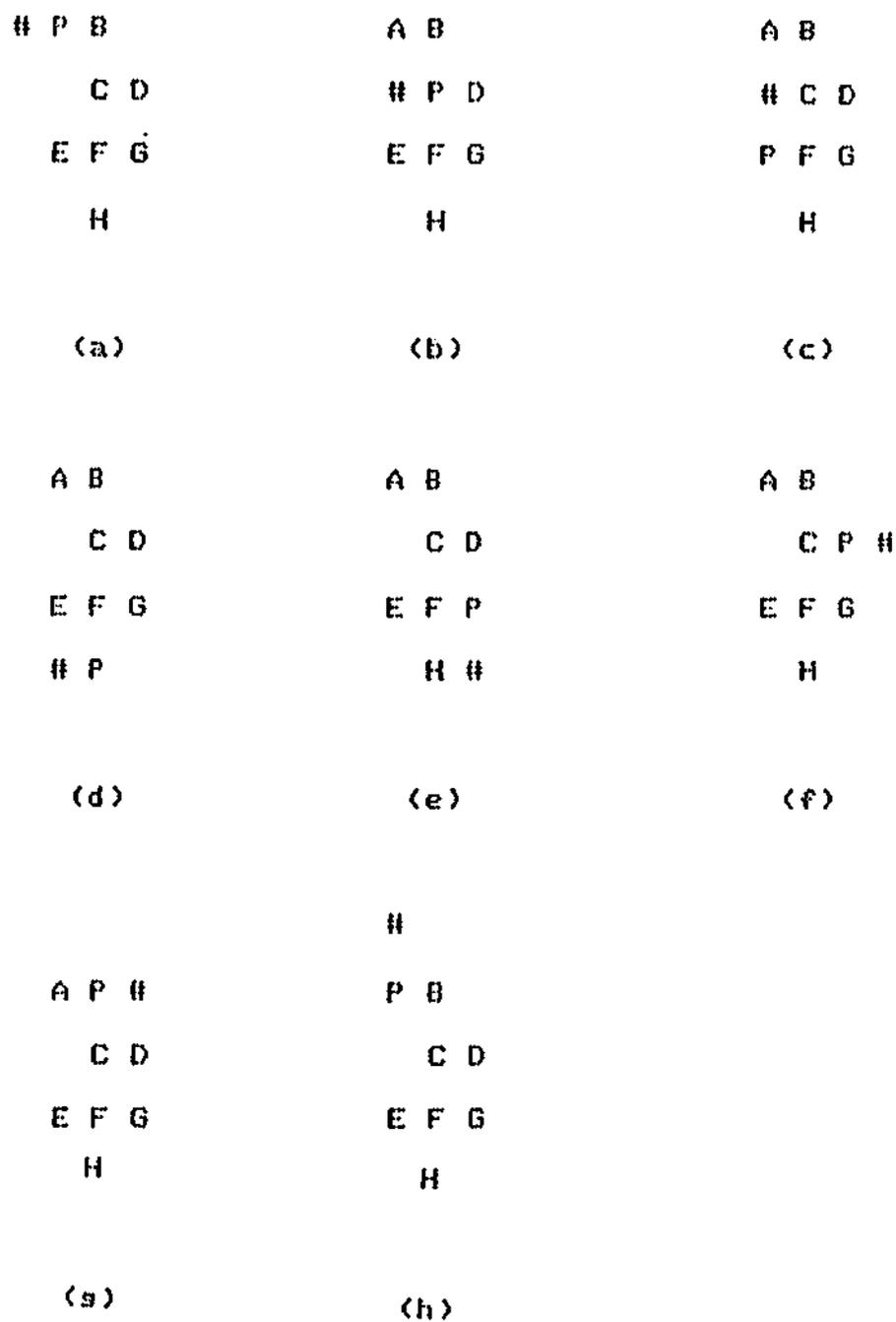


Fig. 9.3 - Percorrimento de borda segundo o Algoritmo 9.1.

A B	A C E H G D B
C D	
E F G	
H	

(a)

(b)

1 B	1 B	1 B
C D	0 1 D	0 1 C
E F G	E F G	1 F G
H	H	H

(c)

(d)

(e)

1 B	1 B	1 B
0 0 1 C	0 0 1 C	0 0 1 1 0
0 1 F G	0 1 F 1	0 1 F 1 0
0 0 1	0 0 1 0	0 0 1 0 0
	0 0 0	0 0 0

(f)

(g)

(h)

	0 0 0	0 0 0 0
1 1 0 0	1 1 0 0	0 1 1 0 0
0 0 1 1 0	0 0 1 1 0	0 0 1 1 0
0 1 F 1 0	0 1 F 1 0	0 1 F 1 0
0 0 1 0 0	0 0 1 0 0	0 0 1 0 0
0 0 0	0 0 0	0 0 0

(i)

(j)

(k)

Fig. 9.4 - Região (a), borda (b), passos na rotulação da borda (c-1).



(a)



(b)



(c)



(d)

Fig. 9.5 - (a) figura original; (b) dois encolhimentos seguidos de duas expansões; (c) duas expansões seguidas de dois encolhimentos; (d) dois encolhimentos.

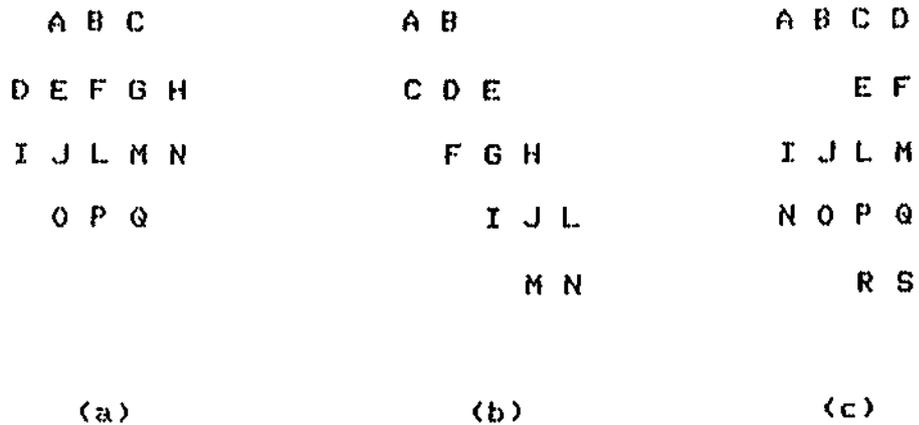


Fig. 9.6 - Regiões com medidas de complexidade $p^2/4$ iguais a (a) 8,70; (b) 11,99; (c) 14,66.

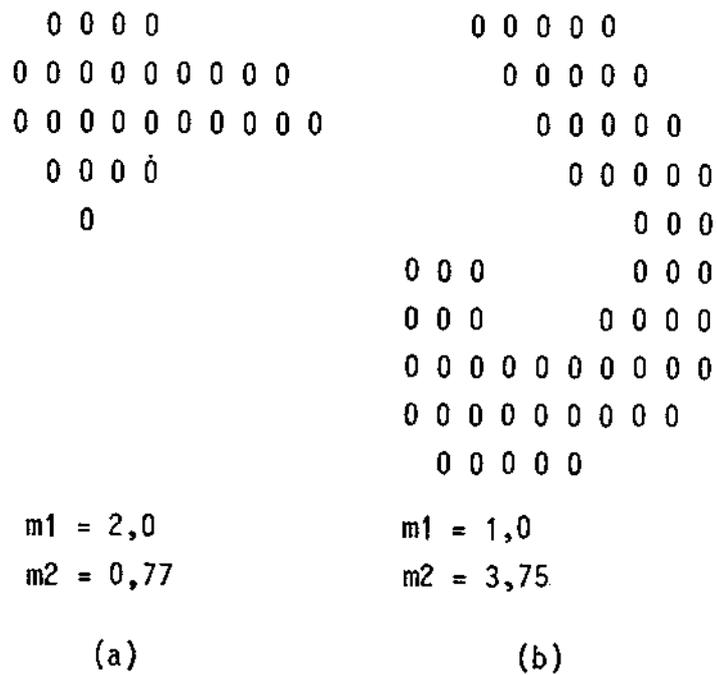


Fig. 9.7 - Medidas de alongamento: $m_1 = \text{m\`a}x\text{ima extens\~ao}/\text{extens\~ao perpendicular}$; $m_2 = \text{\`a}rea/t^2$.

CAPÍTULO 10

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. Assim sendo, têm havido tentativas sérias de automatizar as tarefas de leitura de caracteres alfanuméricos, análise de imagens aeroespaciais, biomédicas e de câmaras de bolha em Física de Alta Energia. Destacam-se também aplicações não-pictóricas como reconhecimento de voz e do orador, análise de sinais de sonar, eletrocardiograma, eletroencefalograma e em sismologia, assim como classificação de dados obtidos em Antropologia ou Zoologia, etc.

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

- a) 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 de terminísticas;
- b) reconhecimento sintático (ou estrutural) de padrões, onde se procura uma descrição de objeto em termos de relações entre suas partes, utilizando principalmente a teoria de linguagens formais.

Este capítulo aborda de maneira resumida apenas as técnicas de classificação de padrões que têm sido mais utilizadas em áreas como detecção de recursos naturais. Para um tratamento aprofundado de reconhecimento sintático de padrões o leitor pode consultar Fu (1974).

10.1 - NOÇÕES BÁSICAS DE CLASSIFICAÇÃO

Os conhecimentos atuais de neurofisiologia não têm sido suficientes para reproduzir, na máquina, as funções 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 entre diversas classes. Os métodos desenvolvidos têm tido por base, em geral, uma formulação estatística que pode estar distante dos processos de percepção em seres vivos.

Deste modo, é feito, inicialmente um conjunto de medidas (usualmente denominadas atributos) 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 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 (denominado espaço de atributos). 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 Figura 10.1 apresenta o esquema geral de um processo de classificação de imagens.

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 caem sempre na partição correspondente. Como esse objetivo ideal nem sempre pode ser alcançado deve-se estabelecer a partição minimizando algum critério; se se utiliza a formulação estatística do problema este critério 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. Essas variações podem ser tanto devidas à incerteza inerente à caracterização da classe, como oriundas de perturbações estatísticas (ruído) nos sensores de medida. A Figura 10.2 ilustra a distribuição de amostras de duas classes (área urbana e mata) obtidas por uma imagem do satélite LANDSAT em dois canais.

10.2 - 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 com vistas à diminuição do espaço computacional requerido. Estas técnicas podem ser divididas em duas categorias: transformações no espaço de atributos ou medidas de informação e distância.

10.2.1 - Transformações no espaço de atributos

O objetivo aqui é transformar o espaço de atributos originais em um espaço de menor dimensionalidade, preservando ao máximo a representação dos padrões e/ou a discriminação entre as classes.

No primeiro caso tem sido usados os critérios de erro médio quadrático ou entropia. A transformação de Karhunen-Loeve (também denominada Transformação de Hotelling ou de principais componentes), já apresentada no capítulo sobre codificação, é também utilizada aqui para efetuar uma redução da dimensionalidade ótima sob o critério de erro médio quadrático. Assim, considerando-se a distribuição multivariada global da mistura das classes e obtendo-se a matriz de covariância \underline{C}_X ($N \times N$) do vetor aleatório \underline{X} ($N \times N$) de atributos, a transformação será dada pela matriz \underline{A} ($P \times N$) tal que as linhas dessa matriz são os vetores próprios correspondentes aos maiores valores de \underline{C}_X (veja o Apêndice). As componentes do vetor $\underline{Y} = \underline{A} \underline{X}$ tornam-se não-correlacionadas. A Figura 10.3 ilustra a Transformação de Karhunen-Loeve para uma distribuição gaussiana bidimensional. Nos novos eixos correlacionados as duas componentes do vetor gaussiano são não-correlacionadas. Isso é evidenciado pelo fato de a elipse que contém uma porcentagem especificada da distribuição gaussiana ter seus eixos principais paralelos aos eixos do novo sistema de coordenadas. As variâncias das componentes (que determinam o alongamento da elipse em cada direção) são os valores próprios da matriz de covariância do vetor aleatório.

A entropia de uma distribuição, definida por:

$$H(\underline{X}) = - E[\log f_{\underline{X}}(\underline{x})], \quad (10.1)$$

pode também ser usada como critério para seleção de atributos, procurando-se, segundo este critério, a Transformação $\underline{Y} = \underline{A}\underline{X}$ que preserva o maior conteúdo possível de informação.

Quando \underline{X} tem uma distribuição gaussiana com matriz de covariância \underline{C} ($N \times N$) é possível mostrar (Young and Calvert, 1974) que:

$$H(\underline{X}) = 1/2 \log |\underline{C}| + N/2 \log 2\pi e, \quad (10.2)$$

onde $|\underline{C}|$ denota o determinante de \underline{C} . Pode-se também mostrar que dentre as transformações ortonormais, aquela que maximiza a entropia de \underline{Y} é dada pela Transformação de Karhunen-Loeve, e a máxima entropia é dada por:

$$H(\underline{Y}) = 1/2 \sum_{i=1}^P \log \lambda_i + P/2 \log 2\pi e, \quad (10.3)$$

onde P é a dimensionalidade do vetor \underline{Y} , e λ_i os autovalores da matriz de transformação.

Quando se procura uma transformação que minimize a separabilidade das classes num espaço de menor dimensionalidade 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, nesse caso, é necessário usar técnicas iterativas ou interativas através de um "display" (Chien, 1978).

Dentre as transformações lineares, deve ser apresentada a função discriminante linear de Fisher: considere-se um problema de classificação de duas classes. Procura-se uma projeção dos dados multivariados em um eixo de tal modo a maximizar a separação entre as classes nesse eixo, isto é, procura-se um vetor \underline{w} tal que (Duda and Hart, 1973)

$$y = \underline{w}^T \underline{x} \quad (10.4)$$

representa a projeção dos dados no eixo procurado.

Sejam

$$\underline{m}_i = 1/n_i \sum_{x \in S_i} \underline{x} \quad i = 1,2 \quad (10.5)$$

as médias amostrais das classes S_i , $i=1,2$ onde n_i é o número de amostras da classe S_i . Sejam as matrizes de espalhamento de cada classe definidas por:

$$\underline{S}_i = \sum_{x \in S_i} (\underline{x} - \underline{m}_i) (\underline{x} - \underline{m}_i)^T \quad i = 1,2. \quad (10.6)$$

Defina-se também a matriz de espalhamento total por:

$$\underline{S}_D = \underline{S}_1 + \underline{S}_2. \quad (10.7)$$

Quando as amostras são projetadas no eixo y , podem ser obtidas as médias amostrais das projeções por:

$$m_i' = 1/n_i \sum_{y \in S_i} y = 1/n_i \sum_{x \in S_i} \underline{w}^T \underline{x} = \underline{w}^T \underline{m}_i \quad i = 1,2 \quad (10.8)$$

Do mesmo modo, o espalhamento das projeções de cada classe é obtido por:

$$s_i^2 = \sum_{y \in S_i} (y - m_i')^2 \quad i = 1,2. \quad (10.9)$$

Portanto $s_1^2 + s_2^2$ representa o espalhamento total das projeções.

Nessas condições, a função discriminante de Fisher maximiza o critério:

$$J(\underline{w}) = \frac{|m_1' - m_2'|^2}{s_1^2 + s_2^2}. \quad (10.10)$$

É possível obter $J(\underline{w})$ como função de \underline{w} . Para tanto, a Equação 10.9 pode ser reescrita como:

$$\begin{aligned} s_i^2 &= \sum_{x \in S_i} (\underline{w}^T \underline{x} - \underline{w}^T \underline{m}_i)^2 = \sum_{x \in S_i} \underline{w}^T (\underline{x} - \underline{m}_i) (\underline{x} - \underline{m}_i)^T \underline{w} = \\ &= \underline{w}^T \underline{S}_i \underline{w}. \end{aligned} \quad (10.11)$$

Daí resulta:

$$\bar{s}_1^2 + \bar{s}_2^2 = \underline{w}^T \underline{S}_D \underline{w}. \quad (10.12)$$

Por outro lado

$$(\underline{m}_1' - \underline{m}_2')^2 = (\underline{w}^T \underline{m}_1 - \underline{w}^T \underline{m}_2)^2 = \underline{w}^T (\underline{m}_1 - \underline{m}_2) (\underline{m}_1 - \underline{m}_2)^T \underline{w} \quad (10.13)$$

Definindo a matriz \underline{S}_E por

$$\underline{S}_E = (\underline{m}_1 - \underline{m}_2) (\underline{m}_1 - \underline{m}_2)^T \quad (10.14)$$

obtem-se

$$(\underline{m}_1' - \underline{m}_2')^2 = \underline{w}^T \underline{S}_E \underline{w}. \quad (10.15)$$

\underline{S}_D é também chamada matriz de espalhamento intraclases, enquanto \underline{S}_E é a matriz de espalhamento entre classes. Deste modo, $J(\underline{w})$ pode ser dado por:

$$J(\underline{w}) = \frac{\underline{w}^T \underline{S}_E \underline{w}}{\underline{w}^T \underline{S}_D \underline{w}}. \quad (10.16)$$

Esta razão é o chamado quociente generalizado de Rayleigh (Duda and Hart, 1973) e pode-se mostrar que sua maximização leva a um problema generalizado de vetor próprio do tipo

$$\underline{S}_E \underline{w} = \lambda \underline{S}_D \underline{w}. \quad (10.17)$$

Supondo-se que \underline{S}_D é não-singular obtêm-se o problema usual de vetor próprio:

$$\underline{S}_D^{-1} \underline{S}_E \underline{w} = \lambda \underline{w}. \quad (10.18)$$

Pela definição de \underline{S}_E verifica-se que o vetor $\underline{S}_E \underline{w}$ está na direção de $(\underline{m}_1 - \underline{m}_2)$, e, como um fator de escala não tem importância no problema obtêm-se a solução:

$$\underline{w} = \underline{S}_D^{-1} (\underline{m}_1 - \underline{m}_2). \quad (10.19)$$

A função discriminante linear de Fisher pode ser generalizada para C classes (Bryan, 1951). Nesse caso a redução de dimensionalidade é feita do espaço original N-dimensional para um espaço (C-1) dimensional (supondo-se $N \geq C$). A dedução da matriz \underline{W} ((C-1)xN) que produz o mapeamento ótimo é análoga ao caso de duas classes, envolvendo as matrizes de espalhamento dentro e entre classes. Para maiores detalhes o leitor pode consultar Duda e Hart (1973).

Deve-se observar que os critérios de representação e os de separabilidade das classes podem ser conflitantes e seria desejável que se estabelecesse um compromisso entre esses fatores na escolha da transformação ótima.

10.2.2 - 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 na redução da dimensionalidade. Assim sendo, utilizam-se critérios indiretos expressos por medidas de distância entre distribuições, os quais fornecem, os quais fornecem limites superiores e/ou inferiores para a probabilidade de erro. Várias medidas têm sido propostas e investigadas as relações entre elas (Kana1, 1974).

São discutidas a seguir duas medidas de separabilidade entre distribuições que têm sido bastante utilizadas (Swain et alii, 1978).

10.2.2.1 - Divergência

Dadas duas classes W_i e W_j , caracterizadas pelas funções de densidade de probabilidade condicionais $f(\underline{x}/W_i)$ e $f(\underline{x}/W_j)$, pode-se definir a razão de verossimilhança $L_{ij}(\underline{x})$ por:

$$L_{ij}(\underline{x}) = \frac{f(\underline{x}/W_i)}{f(\underline{x}/W_j)}, \quad (10.20)$$

assim como o logaritmo dessa razão, ou seja,

$$L_{ij}(\underline{x}) = \log_e L_{ij}(\underline{x}) = \log_e f(\underline{x}/W_i) - \log_e f(\underline{x}/W_j). \quad (10.21)$$

A divergência entre as classes W_i e W_j é definida por:

$$D_{ij} = E[L_{ij}'(\underline{X})/W_i] + E[L_{ji}'(\underline{X})/W_j], \quad (10.22)$$

onde

$$E[L_{ij}'(\underline{X})/W_i] = \int_{\underline{x}} L_{ij}'(\underline{x}) f(\underline{x}/W_i) d\underline{x}, \quad (10.23)$$

$$E[L_{ji}'(\underline{X})/W_j] = \int_{\underline{x}} L_{ji}'(\underline{x}) f(\underline{x}/W_j) d\underline{x}. \quad (10.24)$$

A expressão 10.23 pode ser interpretada com a informação média para discriminação em favor de W_i contra W_j . Uma interpretação análoga é válida para a Equação 10.24 trocando-se as papéis de W_i e W_j . Assim sendo, D_{ij} é uma medida da informação para a discriminação das duas classes.

Pode-se mostrar que as seguintes propriedades são válidas:

a) $D_{ij} > 0$ para $i \neq j$,

b) $D_{ii} = 0$,

c) $D_{ij} = D_{ji}$,

d) Se os atributos x_1, x_2, \dots, x_n são independentes

$$D_{ij}(x_1, x_2, \dots, x_n) = \sum_{k=1}^n D_{ij}(x_k),$$

e) $D_{ij}(x_1, x_2, \dots, x_n, x_{n+1}) \geq D_{ij}(x_1, x_2, \dots, x_n)$.

No caso das duas classes descritas por distribuições gaussianas multivariadas, isto é, $f(\underline{x}/W_i) \sim N(\underline{\mu}_i, \underline{\Sigma}_i)$ e $f(\underline{x}/W_j) \sim N(\underline{\mu}_j, \underline{\Sigma}_j)$, a divergência tem a seguinte expressão:

$$D_{ij} = 1/2 \operatorname{tr}[(\underline{\Sigma}_i - \underline{\Sigma}_j) (\underline{\Sigma}_j^{-1} - \underline{\Sigma}_i^{-1})] + \\ + 1/2 \operatorname{tr}[(\underline{\Sigma}_i^{-1} + \underline{\Sigma}_j^{-1}) (\underline{\mu}_i - \underline{\mu}_j) (\underline{\mu}_i - \underline{\mu}_j)^T]. \quad (10.25)$$

Se as matrizes da covariância $\underline{\Sigma}_i$ e $\underline{\Sigma}_j$ são iguais (i.e., $\underline{\Sigma}_i = \underline{\Sigma}_j = \underline{\Sigma}$) obtêm-se

$$D_{ij} = (\underline{\mu}_i - \underline{\mu}_j)^T \underline{\Sigma}^{-1} (\underline{\mu}_i - \underline{\mu}_j). \quad (10.26)$$

Deve-se notar que o primeiro termo da Expressão 10.25 é devido à diferença entre as matrizes de covariância, enquanto o segundo termo expressa uma distância normalizada entre as médias.

Dadas duas classes a seleção de atributos pela divergência pode ser feita para uma dada redução de dimensionalidade (de, diga-se N para

P) computando a divergência para todos os possíveis $\binom{N}{P}$ subconjuntos de atributos selecionando aquele para o qual a divergência é máxima. O problema que pode ocorrer é o esforço computacional para grandes valores de N e P.

Quando se faz a redução de dimensionalidade na classe mais ampla das transformações lineares ortonormais é possível obter a transformação ótima no caso gaussiano. Distinguem-se 3 casos:

a) $\underline{\Sigma}_i = \underline{\Sigma}_j = \underline{\Sigma}$

É possível mostrar (veja-se Young and Calvert, (1974) ou Fukunaga (1972)) que a discriminação baseada apenas em um atributo dado por:

$$y = (\underline{\mu}_i - \underline{\mu}_j)^T \underline{\Sigma}^{-1} \underline{x} \quad (10.27)$$

não reduz a divergência em relação à situação original de N atributos (x_1, x_2, \dots, x_N) .

b) $\underline{\mu}_i = \underline{\mu}_j$

Nesse caso os atributos ótimos são dados pelos produtos dos vetores próprios de $\underline{\Sigma}_i^{-1} \underline{\Sigma}_j$ com \underline{x} . Os atributos mais significativos são selecionados calculando os correspondentes valores próprios na seguinte ordem:

$$\lambda_1 + 1/\lambda_1 > \lambda_2 + 1/\lambda_2 > \dots > \lambda_N + 1/\lambda_N \quad (10.28)$$

c) Caso geral

Não se conhece procedimento analítico, havendo a necessidade de técnicas de busca para achar a matriz de transformação numericamente (Tou and Heydorn, 1967).

O conceito de divergência é definido para duas classes, e uma generalização para $M > 2$ classes ainda não foi feita. Um procedimento usual é computar a divergência média para todos os pares de classes e selecionar o subconjunto de N atributos para o qual a divergência média, definida por:

$$D_m = \frac{2}{M(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^M D_{ij}, \quad (10.29)$$

é máxima. Todavia, essa estratégia não é ótima. Por exemplo, um único termo na somatória da Expressão 10.29 pode aumentar substancialmente a divergência média. Portanto, no processo de seleção de atributos, é útil examinar cada divergência entre classes (Moik,1980).

10.2.2.2 - Distância J-M

O comportamento da divergência D_{ij} em relação à probabilidade de classificação não é consistente. Enquanto D_{ij} cresce sem limite quando a separabilidade das classes aumenta, aquela probabilidade tende a saturar no valor unitário. Esse comportamento tende a agravar o problema citado no último parágrafo da seção anterior.

Uma medida de separabilidade que evita esse problema é a Distância de Jeffries-Matusita (Distância J-M) (Matusita, 1965), definida por:

$$J_{ij} = \left\{ \int \left[\sqrt{f(x/W_i)} - \sqrt{f(x/W_j)} \right]^2 dx \right\}^{1/2}. \quad (10.30)$$

No caso das densidades gaussianas a Equação 10.30 assume a forma:

$$J_{ij} = [2(1 - e^{-\alpha})]^{1/2}, \quad (10.31)$$

onde

$$\alpha = 1/8 \left(\underline{\mu}_i - \underline{\mu}_j \right)^T \left(\frac{\underline{\Sigma}_i + \underline{\Sigma}_j}{2} \right)^{-1/2} \left(\underline{\mu}_i - \underline{\mu}_j \right) +$$

$$+ 1/2 \log_e \left[\frac{\left| \frac{(\underline{\Sigma}_i + \underline{\Sigma}_j)}{2} \right|}{\left(|\underline{\Sigma}_i| \cdot |\underline{\Sigma}_j| \right)^{1/2}} \right]. \quad (10.32)$$

O comportamento saturante da distância J-M devido ao termo exponencial na Expressão 10.31 tende a reduzir o efeito de classes muito separadas quando se toma a distância J-M média por uma somatória análoga àquela dada pela Equação 10.29.

Uma outra função também utilizada para seleção de atributos que envolve múltiplas classes, com características saturantes análogas à distância J-M, é a divergência transformada, definida por (Swain et alii, 1978).

$$D_{ij}^T = 2 \left[1 - e^{-\frac{D_{ij}}{8}} \right]. \quad (10.33)$$

Deve-se observar neste ponto que a distância J-M está relacionada a outra medida bastante utilizada para seleção de atributos, a distância de Bhattacharyya, definida por:

$$B_{ij} = -\log \rho_{ij}, \quad (10.34)$$

onde

$$\rho_{ij} = \int_{\underline{x}} \sqrt{f(\underline{x}/W_i) f(\underline{x}/W_j)} \, d\underline{x}. \quad (10.35)$$

é o chamado coeficiente de Bhattacharyya (Bhattacharyya, 1943).

É fácil verificar que

$$J_{ij}^2 = 2(1 - \rho_{ij}) \quad (10.36)$$

e que, no caso gaussiano, B_{ij} é dado pela expressão 10.32.

10.3 - ELEMENTOS DE TEORIA ESTATÍSTICA DE DECISÃO

A classificação estatística de imagens é formalizada pela teoria estatística de decisão. Segundo tal teoria, os atributos x_i , $i = 1, 2, \dots, N$ são variáveis aleatórias, que constituem o vetor aleatório \underline{x} . Supõe-se que a função densidade de probabilidade associada a cada classe W_i , $f(\underline{x}/W_i)$, $i = 1, 2, \dots, M$ seja conhecida, assim como as probabilidades a priori $P(W_i)$, $i = 1, 2, \dots, M$ de cada classe. A chamada formulação bayesiana envolve também custos $C(W_k/W_i)$, $k, i = 1, 2, \dots, M$ de tomar uma decisão por uma classe W_k quando a verdadeira classe é W_i .

Deste modo, sendo observado o vetor de atributos \underline{x} , o risco médio que incorre ao tomar a decisão W_k é dado por:

$$L_{\underline{x}}(W_k) = \sum_{i=1}^M C(W_k/W_i) P(W_i|\underline{x}). \quad (10.37)$$

Deve-se tomar a decisão W_k que minimize o risco médio sobre a distribuição de \underline{x} , ou seja,

$$L(W_k) = \int_{\underline{x}} L_{\underline{x}}(W_k) f(\underline{x}) d\underline{x}. \quad (10.38)$$

Torna-se claro que $L(W_k)$ é minimizado se, para cada valor de \underline{x} , $L_{\underline{x}}(W_k)$ seja minimizado.

Um caso particular da formulação bayesiana de importância é obtido quando se tomam os custos $C(W_k|W_i)$ na forma

$$\begin{aligned} C(W_k|W_i) &= 0 \quad i = k, \\ &= 1 \quad i \neq k \quad i, k = 1, 2, \dots, M. \end{aligned} \tag{10.39}$$

Nessas condições, a Expressão 10.37 assume a forma:

$$L_{\underline{x}}(W_k) = 1 - P(W_k|\underline{x}) \tag{10.40}$$

e a minimização de $L_{\underline{x}}(W_k)$ é feita se, para cada vetor \underline{x} de atributos observado, decidir-se pela classe W_k que maximize a probabilidade a posteriori $P(W_k|\underline{x})$.

Utilizando a regra de Bayes tem-se:

$$P(W_k|\underline{x}) = \frac{f(\underline{x}|W_k) P(W_k)}{f(\underline{x})} \tag{10.41}$$

e, como o denominador é independente de k , pode-se atribuir o vetor \underline{x} à classe k que maximize a expressão $f(\underline{x}|W_k) P(W_k)$.

Pode-se, dessa maneira, encarar o processo de classificação como dado pelo cômputo de funções discriminantes:

$$g_i(\underline{x}) = f(\underline{x}|W_i) P(W_i) \quad i = 1, 2, \dots, M. \tag{10.42}$$

para cada vetor de atributos \underline{x} observado e pela distribuição desse vetor à classe W_k que forneça a máxima função discriminante. Tal esquema está ilustrado na Figura 10.4.

É frequente a situação em que se admitem as probabilidades a priori $P(W_i)$, $i = 1, 2, \dots, M$ iguais a $1/M$. Nesse caso as funções discrimi

nantes se resumem às funções densidade de probabilidade condicionais $f(\underline{x}|W_i)$. Tal esquema de decisão é denominado decisão por máxima verossimilhança.

É possível tomar a decisão computando as funções discriminantes dadas pela Expressão 10.42 ou ainda uma função monotonicamente crescente de $g_i(\underline{x})$, por exemplo

$$\log_e g_i(\underline{x}) = \log_e f(\underline{x}|W_i) + \log_e P(W_i). \quad (10.43)$$

A implementação do classificador pela Equação 10.42 ou 10.43 equivale a uma partição do espaço N-dimensional dos vetores de atributos \underline{x} em regiões D_i , $i = 1, 2, \dots, M$. A região D_i , atribuída à classe W_i , é aquela em que:

$$g_i(\underline{x}) > g_k(\underline{x}) \quad i \neq k. \quad (10.44)$$

No caso de duas classes o cálculo da probabilidade de erro de decisão ($P(\epsilon)$) pode ser feito observando que há duas maneiras de cometer um erro, isto é, o vetor \underline{x} provém da classe W_1 e decidiu-se por W_2 ou vice-versa. Deste modo, tem-se (Duda and Hart, 1973).

$$\begin{aligned} P(\epsilon) &= P(\underline{x} \in D_2, W_1) + P(\underline{x} \in D_1, W_2) \\ &= P(\underline{x} \in D_2 / W_1) P(W_1) + P(\underline{x} \in D_1 / W_2) P(W_2) \\ &= \int_{R_2} f(\underline{x}|W_1) P(W_1) d\underline{x} + \int_{R_1} f(\underline{x}|W_2) P(W_2) d\underline{x} \end{aligned} \quad (10.45)$$

A Figura 10.5 ilustra o cálculo de $P(\epsilon)$ no caso de duas distribuições gaussianas unidimensionais.

No caso de múltiplas classes torna-se mais simples calcular a probabilidade de classificação correta ($P(C)$), que é dada por:

$$\begin{aligned} P(C) &= \sum_{i=1}^M P(\underline{x} \in D_i, W_i) \\ &= \sum_{i=1}^M P(\underline{x} \in D_i/W_i) \cdot P(W_i) \\ &= \sum_{i=1}^M \int_{D_i} f(\underline{x}/W_i) P(W_i) d\underline{x} \end{aligned} \quad (10.46)$$

Deve-se observar que, embora formalmente resolvido, o problema de cálculo da probabilidade de erro pode tornar-se difícil na situação multivariada.

Densidades multivariadas gaussianas têm sido frequentemente utilizadas para caracterizar as classes, particularmente na área de Sensoriamento Remoto (Crane et alii, 1972). A forma geral dessa densidade é dada por:

$$f(\underline{x}/W_i) = \frac{1}{(2\pi)^{N/2} |\underline{C}_i|^{1/2}} e^{-1/2(\underline{x}-\underline{\mu}_i)^T \underline{C}_i^{-1}(\underline{x}-\underline{\mu}_i)}, \quad (10.47)$$

onde \underline{x} é o vetor (Nx1) de atributos, $\underline{\mu}_i$ é o vetor (Nx1) de valores esperados, \underline{C}_i (NxN) é a matriz de covariância e $|\underline{C}_i|$ denota o determinante dessa matriz.

A Figura 10.6 ilustra as curvas de nível de uma densidade gaussiana bidimensional.

No caso multidimensional essas curvas de nível são elipsóides representados pela equação:

$$d = (\underline{x}-\underline{\mu}_i)^T \underline{C}_i^{-1}(\underline{x}-\underline{\mu}_i). \quad (10.48)$$

A quantidade d representa a chamada distância de Mahalanobis de \underline{x} a $\underline{\mu}_i$. Os eixos principais do elipsóide são os vetores próprios de \underline{C}_i e os valores próprios determinam o comprimento desses eixos.

Utilizando a Expressão 10.47 na 10.43 e notando que o termo $-N/2 \log_e 2\pi$ não depende de i , obtêm-se as funções discriminantes para o caso gaussiano multivariado:

$$g_i^1(\underline{x}) = -1/2 \log |\underline{C}_i| - 1/2 (\underline{x} - \underline{\mu}_i)^T \underline{C}_i^{-1} (\underline{x} - \underline{\mu}_i) + \log P(W_i). \quad (10.49)$$

A Figura 10.7 mostra uma imagem LANDSAT (canal 7) obtida sobre a região de São José dos Campos. A classificação dessa imagem foi feita utilizando a Equação 10.49, supondo as classes equiprováveis, e o resultado da classificação que evidencia a classe denominada "área urbana" está mostrado na Figura 10.8.

O esquema de classificação proposto até este ponto atribui um vetor de atributos \underline{x} a alguma classe entre as M prováveis. Em certas circunstâncias, porém, o valor de $g_i^1(\underline{x})$ (ou $g_i^1(\underline{x})$) é suficientemente baixo para todo $i = 1, 2, \dots, M$ a tal modo que seria natural rejeitar o padrão ao invés de atribuí-lo a alguma classe. Isso pode ser feito observando que a forma quadrática $(\underline{x} - \underline{\mu}_i)^T \underline{C}_i^{-1} (\underline{x} - \underline{\mu}_i)$ tem a distribuição qui-quadrado com N graus de liberdade. Portanto, utilizando a Expressão 10.49, pode-se computar a expressão:

$$Q_i(\underline{x}) = -2 g_i^1(\underline{x}) - \log_e |\underline{C}_i| + 2 \log_e P(W_i). \quad (10.50)$$

Se o valor de $Q_i(\underline{x})$ ultrapassar o limiar da distribuição qui-quadrado correspondente a uma dada porcentagem acumulada da distribuição, rejeita-se o vetor de atributos.

A implementação direta do classificador gaussiano requer o computador da função discriminante dado pela Expressão 10.49 para cada vetor de atributos. O tempo de classificação é proporcional a $N^2.M$, onde N é a dimen

são do vetor de atributos e M é o número de classes. Em algumas situações, como em Sensoriamento Remoto, o número de vetores de atributos realmente existentes em uma imagem é suficientemente limitado para que os esquemas do tipo tabela por espalhamento ("hashing") possam ser empregados (Shlien, 1975), abreviando consideravelmente o esforço computacional.

Com a mesma finalidade, têm também sido utilizados os chamados classificadores de paralelepípedos, em que cada classe é definida pelos limites superior e inferior de um histograma obtido numa área de treinamento (área de classificação conhecida). 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 probabilidade de utilizar os esquemas de testes sequenciais de hipóteses tem sido explorada em classificação de padrões, dado o custo substancial por vezes inerente à aquisição de cada atributo. Fu (1968) apresenta vários resultados da aplicação desta técnica na classificação de caracteres alfanuméricos manuscritos.

10.4 - TREINAMENTO E APRENDIZADO

Na seção anterior o problema de classificação foi resolvido sob a hipótese de que as densidades de probabilidade que caracterizam as classes são conhecidas. Essa hipótese não é frequentemente satisfeita num problema de reconhecimento de padrões e os métodos de classificação descritos abreviadamente a seguir são utilizados quando a hipótese mencionada anteriormente não é satisfeita.

10.4.1 - Aprendizado com supervisão

Quando a forma funcional das densidades é conhecida, exceto por seus parâmetros, tem-se o chamado problema de aprendizado, que é um processo de estimação dos parâmetros desconhecidos. Quando existem amostras disponíveis de classificação conhecida, diz-se que o problema de aprendizado é supervisionado.

Em imagens de sensoriamento remoto, por exemplo, a hipótese gaussiana é frequentemente utilizada para caracterizar as densidades multivariadas dos elementos das imagens multiespectrais. Todavia, os vetores média e as matrizes de covariância de cada classe em geral não são conhecidos a priori.

Para resolver esse problema, dois modelos podem ser adotados: se os parâmetros a serem estimados 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.

Em sensoriamento remoto, o primeiro modelo é geralmente adotado e, a partir das amostras das áreas de classificação conhecida, podem ser estimados o vetor de valores esperados da classe W_i por:

$$\hat{\underline{\mu}}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} \underline{x}_j \quad (10.51)$$

e a matriz de covariância da classe W_i por:

$$\hat{\underline{C}}_i = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} \left(\underline{x}_j - \hat{\underline{\mu}}_i \right) \left(\underline{x}_j - \hat{\underline{\mu}}_i \right)^T, \quad (10.52)$$

onde M_i é o número de amostras disponíveis na área de treinamento da classe W_i .

Esses estimadores podem ser expressos em forma recursiva (Tou and Gonzalez, 1974). Denota-se o estimador de $\underline{\mu}_i$ baseado em N_i amostras por $\hat{\underline{\mu}}_i(N_i)$. O estimador de $\underline{\mu}_i$ baseado em N_{i+1} amostras pode ser dado por:

$$\hat{\underline{\mu}}_i(N_{i+1}) = \frac{1}{N_{i+1}} \left(N_i \hat{\underline{\mu}}_i(N_i) + \underline{x}_{N_{i+1}} \right). \quad (10.53)$$

O processo de recursão inicia-se com $\hat{\underline{\mu}}_i(1) = \underline{x}_1$.

Analogamente, para o estimador da matriz de covariância, tem-se:

$$\begin{aligned} \hat{\underline{C}}_i(N_{i+1}) = & \frac{1}{N_{i+1}} \left\{ N_i \hat{\underline{C}}_i(N_i) + N_i \hat{\underline{\mu}}_i(N_i) \hat{\underline{\mu}}_i^T(N_i) + \right. \\ & \left. + \frac{\underline{x}_{N_{i+1}} \underline{x}_{N_{i+1}}^T}{N_{i+1}} - \frac{1}{(N_{i+1})^2} \left\{ N_i \hat{\underline{\mu}}_i(N_i) + \underline{x}_{N_{i+1}} \right\} \cdot \right. \\ & \left. \cdot \left\{ N_i \hat{\underline{\mu}}_i(N_i) + \underline{x}_{N_{i+1}} \right\}^T \right\}. \end{aligned} \quad (10.54)$$

O processo de recursão inicia-se com:

$$\hat{\underline{C}}_i(1) = \underline{x}_1 \underline{x}_1^T - \hat{\underline{\mu}}_i(1) \cdot \hat{\underline{\mu}}_i^T(1) = \underline{0}. \quad (10.55)$$

10.4.2 - Aprendizado sem supervisão

No aprendizado sem supervisão as amostras na área de treinamento não são rotuladas. A densidade de probabilidade desses vetores é a densidade-mistura, dada por:

$$f(\underline{x}) = \sum_{i=1}^M f(\underline{x}|W_i) P(W_i). \quad (10.56)$$

A solução do problema existe se o número de classes (M), e as probabilidades a priori ($P(W_i)$) são conhecidas, assim como a forma das densidades $f(\underline{x}|W_i)$, exceto por seus parâmetros (Duda and Hart, 1973).

10.4.3 - Técnicas não-paramétricas

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, utilizando as técnicas não-paramétricas troca-se a necessidade da forma funcional das densidades ser conhecida pela necessidade de um número maior de amostras.

10.4.4 - Algoritmos de treinamento

Em várias situações, a única informação oferecida ao sistema de classificação consiste em amostras de classificação conhecida. Nessas condições, o classificador deve ajustar iterativamente as superfícies de decisão de modo a, se possível, separar perfeitamente as amostras.

O caso mais simples que se apresenta é o de duas classes cujas amostras podem ser separadas no espaço euclidiano N-dimensional por um hiperplano. Essa superfície pode ser caracterizada pela equação $g(\underline{x}) = 0$, onde:

$$g(\underline{x}) = p_0 + \sum_{i=1}^N p_i x_i. \quad (10.57)$$

Nas amostras da classe W_1 deve-se ter $g(\underline{x}) > 0$, e o oposto deve ocorrer para amostras da classe W_2 .

A Expressão 10.57 pode ser colocada numa forma homogênea se forem definidos vetores aumentados de uma dimensão por meio de:

$$\underline{y} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \quad \text{e} \quad \underline{a} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \\ p_N \end{pmatrix}. \quad (10.58)$$

Nessas condições a determinação do hiperplano de separação consiste em achar um vetor \underline{a} , tal que:

$$\underline{a}^T \underline{y}_1^j > 0 \quad \text{para} \quad j = 1, 2, \dots, M_1 \quad (10.59)$$

$$\text{e} \quad \underline{a}^T \underline{y}_2^j < 0 \quad \text{para} \quad j = 1, 2, \dots, M_2, \quad (10.60)$$

onde \underline{y}_1^j , $j = 1, 2, \dots, M_1$ são as amostras da classe W_1 e \underline{y}_2^j , $j = 1, 2, \dots$, são as amostras da classe W_2 . Se as amostras da classe W_2 forem multiplicadas por (-1) , o problema se resume em encontrar um vetor \underline{a} , tal que:

$$\underline{a}^T \underline{y}^j > 0 \quad j = 1, 2, \dots, M_1 + M_2. \quad (10.61)$$

A Figura 10.9 ilustra esse problema no caso bidimensional. De modo geral, a solução não é única, mas tem-se uma região de solução no espaço euclidiano N-dimensional.

Deste modo, o problema consiste em encontrar uma solução para um sistema de desigualdades lineares. Conforme mencionado anteriormente, tal problema ocorre também nas áreas de restauração e reconstrução de imagens.

Os algoritmos iterativos originalmente propostos para resolver esse problema, no contexto de classificação de padrões, inspiraram-se em modelos simplificados de neurônios baseados na idéia de punição e premiação (Rosenblatt, 1961). Tais algoritmos são genericamente classificados como

algoritmos do tipo "perceptron". A formulação mais simples consiste no seguinte esquema iterativo. Um vetor inicial arbitrário $\underline{a}(0)$ é escolhido. Os protótipos (isto é, amostras de classificação conhecida) são examinados ciclicamente. Se num dado passo

$$\underline{a}^T(j) \underline{y}^j \leq 0, \quad (10.62)$$

então:

$$\underline{a}(j+1) = \underline{a}(j) + \underline{y}^j. \quad (10.63)$$

Caso contrário,

$$\underline{a}(j+1) = \underline{a}(j). \quad (10.64)$$

Pode-se mostrar (Duda and Hart, 1973) que se as classes W_1 e W_2 são linearmente separáveis, o algoritmo converge num número finito (mas desconhecido) de passos.

Várias generalizações do "perceptron" têm sido propostas, incluindo-se o uso de uma margem b de tal modo que uma correção no vetor $\underline{a}(j)$ é feita apenas se $\underline{a}^T(j) \underline{y}(j) \leq b$.

O algoritmo do "perceptron" e seus derivados podem ser formulados como algoritmos do tipo gradiente, em que na região de solução a função escolhida possui um mínimo. No caso da versão mais simples do "perceptron" apresentada, essa função tem a forma:

$$J(\underline{a}, \underline{y}) = 1/2 (|\underline{a}^T \underline{y}| - \underline{a}^T \underline{y}). \quad (10.65)$$

O problema que ocorre com o algoritmo do "perceptron" e suas variações é a não-convergência no caso em que as classes não são separáveis linearmente. Para contornar tal problema, Ho e Kashyap (1965) propuseram um algoritmo que minimiza uma função do tipo erro quadrático, que converge ainda na situação não-separável. Agrupando os $(M_1 + M_2)$ vetores \underline{y} numa matriz

\underline{Y} $((M_1 + M_2) \times (N+1))$ e considerando um vetor de margem \underline{b} $((M_1 + M_2) \times 1)$ todas as componentes positivas, o problema consiste em encontrar um vetor \underline{a} tal que:

$$\underline{Y} \underline{a} \geq \underline{b} > \underline{0}. \quad (10.66)$$

Define-se uma função objetiva quadrática a ser minimizada por:

$$J(\underline{a}, \underline{b}) = 1/2 \|\underline{Y} \underline{a} - \underline{b}\|^2 \quad (10.67)$$

sob a restrição $\underline{b} > \underline{0}$.

O algoritmo iterativo do gradiente (aplicado ciclicamente nos protótipos) assume a forma:

$$\underline{b}(0) > \underline{0} \quad \text{arbitrário}$$

$$\underline{a}(0) = (\underline{Y}^T \underline{Y})^{-1} \underline{Y}^T \underline{b}(0) = \underline{Y}^+ \underline{b}(0)$$

$$\underline{b}(j+1) = \underline{b}(j) + \rho [\underline{e}(j) + |\underline{e}(j)| \underline{j}]$$

$$\underline{a}(j+1) = (\underline{Y}^T \underline{Y})^{-1} \underline{Y}^T \underline{b}(j+1) = \underline{Y}^+ \underline{b}(j+1) \quad (10.68)$$

onde o vetor de erro \underline{e} é dado por:

$$\underline{e}(j) = \underline{Y} \underline{a}(j) - \underline{b}(j) \quad (10.60)$$

Se as amostras de treinamento são linearmente separáveis, a solução das desigualdades lineares é obtida num número finito (mas desconhecido) de passos. Se as amostras são não-separáveis, há evidência de não-separabilidade pela convergência de $\|\underline{e}\|$ para um valor não-nulo num número finito ou infinito de passos.

Na situação de múltiplas classes (M), três situações podem se apresentar:

- a) Uma classe é separável linearmente de todas as outras. Há portanto M decisões binárias a serem tomadas.
- b) Cada classe é separável linearmente de cada outra classe. Há portanto $\binom{M}{2}$ decisões binárias a serem tomadas.
- c) Existem M funções discriminantes lineares $g_i(\underline{x})$ que:

$$g_i(\underline{x}) > g_j(\underline{x}) \text{ para todo } j \neq i.$$

Nesse último caso é possível generalizar o algoritmo de "perceptron" para M classes e provar sua convergência de modo análogo ao caso binário.

Deve-se notar neste ponto que o problema de obter a solução de um sistema de desigualdades lineares pode ser resolvido por algoritmos de programação linear (Charnes et alii, 1953).

Quando as classes não são separáveis por hiperplanos, é possível utilizar os esquemas iterativos de aprendizado do tipo "perceptron" e seus derivados pela discriminação através das chamadas funções discriminantes lineares generalizadas do tipo:

$$g(\underline{x}) = \sum_{i=1}^d a_i y_i(\underline{x}), \quad (10.70)$$

onde $y_i(\underline{x})$ são funções arbitrárias de \underline{x} , e o número de termos d é, em geral, superior à dimensionalidade (N) do vetor de padrões \underline{x} . Tal é o caso de funções do tipo polinômio, cujo grau pode ser selecionado para prover a separação entre classes.

10.4.5 - Técnicas de Agregação de Dados

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 taxono

mia 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. O problema básico de agregação consiste em particionar as nuvens de amostras no espaço N-dimensional segundo esse critério.

Entre as medidas de similaridade propostas, podem ser mencionadas:

- 1) distância euclídeana. Trata-se de uma medida que é invariante a translações, mas não é invariante a transformações lineares em geral.
- 2) Cosseno do ângulo entre dois vetores de atributos \underline{x}_i e \underline{x}_j , que pode ser definido por:

$$S(\underline{x}_i, \underline{x}_j) = \frac{\underline{x}_i^T \underline{x}_j}{\|\underline{x}_i\| \cdot \|\underline{x}_j\|} \quad (10.70)$$

Esta medida é invariante à rotação e mudança de escala, mas não é invariante à translação e a transformações lineares em geral.

- 3) Medida de Tanimoto, definida para atributos binários, e expressa por:

$$T(\underline{x}_i, \underline{x}_j) = \frac{\underline{x}_i^T \underline{x}_j}{\underline{x}_i^T \underline{x}_i + \underline{x}_j^T \underline{x}_j - \underline{x}_i^T \underline{x}_j} \quad (10.71)$$

Uma interpretação da Expressão 10.71 revela que essa medida representa a razão entre o número de atributos comuns a \underline{x}_i e \underline{x}_j e o número de atributos pos

suídos por \underline{x}_i e \underline{x}_j (definindo que \underline{x}_i possui o k ésimo atributo se o k ésimo componente de \underline{x}_i é 1).

Dentre os critérios a serem minimizados pela partição das amostras o mais utilizado envolve a soma dos erros quadráticos, pela expressão:

$$J = \sum_{i=1}^M \sum_{x \in S_i} \|\underline{x} - \underline{m}_i\|^2, \quad (10.72)$$

onde M é o número de classes, $(x \in S_i)$ indica as amostras pertencentes à partição S_i e \underline{m}_i é a média amostral da partição S_i , definida por:

$$\underline{m}_i = \frac{1}{M_i} \sum_{x \in S_i} \underline{x}, \quad (10.73)$$

onde M_i é o número de amostras no agregamento S_i .

Uma partição ótima é definida como aquela que minimiza o critério J dado pela Expressão 10.72. Um algoritmo de agregação típico escolhe uma partição inicial e consiste num procedimento iterativo que reatribui vetores de padrão aos agregamentos até obter um mínimo de J . Em geral obtém-se um mínimo local de J , não havendo garantia de um mínimo global. Diferentes escolhas da partição inicial podem levar a diferentes mínimos locais. Um procedimento iterativo básico para obter a minimização de J é o chamado algoritmo de K médias, descrito a seguir:

- a) Selecione o número de aglomerados K .
- b) Escolha uma partição do número total de amostras (M_t) em K aglomerados. Compute as médias amostrais \underline{m}_i , $i=1,2, \dots, K$ de cada aglomerado.
- c) Calcule a distância do vetor de padrões \underline{x} a cada média \underline{m}_i , $i=1,2, \dots, K$.

- d) Atribua \underline{x} ao aglomerado cuja média esteja mais próxima de \underline{x} .
- e) Se nenhuma reatribuição de amostras a aglomerados ocorreu (ou se o número máximo permitido de iterações foi ultrapassado), pare.
- f) Caso contrário, calcule as novas médias amostrais dos aglomerados e vá para c).

As propriedades de convergência do algoritmo de K médias foram estudadas por Mac Queen (1967). O algoritmo denominado Isodata (Ball and Hall, 1965) representa um desenvolvimento 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 de imagens, tanto no caso escalar (Velasco, 1980) (sendo equivalente ao procedimento de Max) como no caso vetorial (Linde et alii, 1980).

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 agregamento, num nível de 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 famílias, etc. Apesar de sua enorme importância prática, os métodos de agregação frequentemente ressentem-se de uma base mais sólida nas decisões de dividir ou aglomerar amostras, ou na escolha do número de aglomerados. Assim sendo, desenvolve-se atualmente um esforço considerável no sentido de validar estatisticamente estes processos.

10.5 - AVALIAÇÃO DE DESEMPENHO DO CLASSIFICADOR

Em última análise, o desempenho do classificador é avaliado pela probabilidade de erro $P(\text{erro})$ de classificação. No caso de múltiplas classes envolvendo a medida de vários atributos, o cálculo direto da probabilidade de erro torna-se difícil: a solução adotada é estimar essa probabilidade a partir do desempenho do classificador. Sabe-se que o procedimento pa

ra avaliar a P(erro) a partir do desempenho do classificador sobre as próprias amostras de treinamento é otimista. O procedimento usual é dividir as amostras de classificação conhecida em áreas de treinamento e de teste, bem como avaliar o classificador sobre essas últimas.

Se a verdadeira probabilidade de erro é p_e , se M amostras são utilizadas para teste e se m dessas amostras são classificadas erroneamente, a distribuição da variável aleatória m é binomial, ou seja:

$$P(m) = \binom{M}{m} p_e^m (1-p_e)^{M-m}. \quad (10.74)$$

Pode-se mostrar (Fukunaga, 1972) que o estimador de máxima verossimilhança de p_e é dada por:

$$\hat{p}_e = \frac{m}{M} : \quad (10.75)$$

A partir da distribuição desse estimador, é possível obter intervalos de confiança para a estimativa de p_e (Highleyman, 1962). A não ser que M seja razoavelmente grande, o valor obtido pela Equação 10.75 deve ser encarado com cuidado. Por exemplo, se o classificador não cometeu erro algum em 50 amostras, o verdadeiro valor de p_e está entre 0% e 8% com probabilidade de 0,95. Para que esse intervalo de confiança esteja entre 0% e 2% seriam necessárias 250 amostras verificadas sem nenhum erro.

Dado um conjunto de amostras de classificação conhecida disponíveis, surge o problema de como dividi-las em amostras de treinamento e de teste. Se muitas amostras forem selecionadas para treinamento e poucas para teste, a avaliação da P(erro) será pobre; se a estratégia oposta for adotada, o projeto do classificador poderá ficar comprometido. 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 resultantes N-1 são usadas para treinamento. De qualquer modo, a solução do problema relativo à repartição de amostras entre treinamento e teste ainda é objeto de pesquisas (Toussaint, 1974).

A disponibilidade de um número finito de amostras para treinamento pode afetar significativamente o desempenho do classificador, tendo sido observado que, nessa situação, o desempenho do classificador não aumenta indefinidamente com o aumento do número de atributos (Kanai and Chandrasekaran, 1971).

10.6 - DESENVOLVIMENTOS RECENTES E TENDÊNCIAS FUTURAS

Para completar este sumário das aplicações de classificação de padrões à análise de imagens, devem ser mencionados alguns desenvolvimentos recentes e tendências futuras, particularmente no caso de imagens de recursos naturais (Landgrebe, 1981). As imagens de sensoriamento remoto são colhidas em geral em várias bandas do espectro e, como consequência, os métodos tradicionais de classificação são baseados em características espectrais. Recentemente, tem havido um grande esforço para desenvolver algoritmos que incorporam atributos espaciais, variações temporais e dados auxiliares. Dentre as técnicas que incorporam atributos espaciais podem ser mencionados: a) filtragem espacial para extrair atributos locais, que são acrescentados aos atributos espectrais (Dondes and Rosenfeld, 1982, Dutra and Mascarenhas, 1984)); b) matrizes de coocorrência, que são histogramas bidimensionais de níveis de cinza, para uma dada distância e uma dada direção; c) algoritmos de partição de cenas, classificando elementos de imagem por grupos, através de processos de anexação de regiões por critérios estatísticos (Kettig and Landgrebe, 1976); d) teoria de decisão composta, que leva em conta a correlação de elementos de imagem adjacentes na classificação (Swain et alii, 1981; Chittineni, 1981); e) métodos de relaxação (Rosenfeld et alii, 1976), que são processos iterativos de atualização das probabilidades a posteriori das classes baseadas na dependência probabilística com elementos de imagem vizinhos.

Os algoritmos que incorporam características temporais utilizam-se da periodicidade da cobertura de uma dada região pelos satélites de recursos naturais. Essas técnicas pressupõem o registro preciso de passagens diferentes. A primeira abordagem utilizada simplesmente adicionava os atributos espectrais de novas passagens aos atributos da primeira passagem. Todavia, sugem problemas devidos à necessidade de um número maior de amostras de

treinamento como decorrência da maior dimensionalidade das amostras, bem como aqueles devidos à não-disponibilidade de dados por cobertura de nuvens. Têm sido estudados métodos que envolvem o estudo de trajetórias temporais espectrais do tipo de terreno (Misra and Wheeler, 1978), particularmente em levantamento de safra, bem como uma abordagem bayesiana num classificador em cascata (Swain, 1978).

A utilização de dados auxiliares pode se basear em informação subjetiva e descritiva como práticas de cultivo ou condições de precipitação pluvial na área, ou em dados quantitativos como perfil topográfico, levando aos chamados modelos digitais de terreno (Horn, 1981); resposta a outros sensores como radar; mapas de classificação de anos prévios; fronteiras políticas; etc. Um exemplo de uso de dados auxiliares na classificação é aquele que foi utilizado no chamado Procedure 1 do LACIE ("Large Area Crop Inventory Experiment"), um extenso programa de levantamento de safras de trigo desenvolvido por várias agências nos EUA. Nesse caso, o analista dispunha, no processo de rotulação de pontos selecionados dos segmentos a ser classificados, de vários dados auxiliares como plotagem de trajetórias temporais em espaços transformados (Kauth and Thomas, 1976). Deve ser mencionada finalmente uma tendência crescente na análise de recursos naturais, que consiste na utilização dos chamados sistemas geográficos de informação, os quais incorporam dados auxiliares que, associados muitas vezes a imagens de satélites, fornecem informação para o gerenciamento de recursos naturais (Bryant and Zobrist, 1976).

REFERENCIAS BIBLIOGRÁFICAS

- BALL, G.H.; HALL, D.J. Isodata, an iterative method of multivariate analysis and pattern classification. In: IFIPS CONGRESS, *Proceedings*, 1965.
- BHATTACHARYYA, A. On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99-109, 1943.
- BRYAN, J.G. The generalized discriminant function: mathematical foundations and computational routine. *Harvard Educational Review*, 21:90-95, 1951.
- BRYANT, N.A.; ZOBRIST, A.L. IBIS: a geographic information system based on digital image processing and raster datatype. In: Symposium on Machine Processing of Remotely Sensed Data, West Lafayette, IN, *Proceedings*, 1976.
- CHARNES, A.; COOPER, W.W.; HENDERSON, A. *An introduction to linear programming*. New York, NY, John Wiley, 1953.
- CHIEN, Y.T. *Interactive pattern recognition*. New York, NY, Marcel Dekker, 1978.
- CHITTINENI, C.B. Utilization of spectral-spatial information in the classification of imagery data. *Computer Graphics and Image Processing*, 16(8):305-340, Aug. 1981.
- CRANE, R.B.; MALILA, W.A.; RICHARDSON, W. Suitability of the normal density assumption likelihood for processing multispectral scanner. *IEEE Transactions on Geoscience Electronics*, GE-10(4):158-165, 1972.
- DONDES, P.A.; ROSENFELD, A. Pixel classification based on gray level and local busyness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4(1):79-84, Jan. 1982.
- DUDA, R.O.; HART, P.E. *Pattern classification and scene analysis*. New York, NY, John Wiley, 1973.

- DUTRA, L.V.; MASCARENHAS, N.D.A. Some experiments with spatial feature extractions methods in multispectral classification. *International Journal of Remote Sensing*, 5(2):303-313, Mar/Apr. 1984.
- FU, K.S. *Sequential methods in pattern recognition and machine learning*. New York, NY, Academic, 1968.
- *Syntactic methods in pattern recognition*. New York, NY, Academic, 1974.
- FUKUNAGA, K. *Introduction to statistical pattern recognition*. New York, NY, Academic, 1972.
- HIGHLEYMAN, W.H. The design and analysis of pattern recognition experiments. *Bell System Technical Journal*, 41:723, mar. 1962.
- HO, Y.C.; KASHYAP, R.L. An algorithm for linear inequalities and its applications. *IEEE Transactions on Electronic Computers*, EC-14:683-688, Oct. 1965.
- HORN, B.K.P. Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1):14-47, Jan. 1981.
- KANAL, L.N. Patterns in pattern recognition 1968-1974. *IEEE Transactions on Information Theory*, IT-20(6):697-722, Nov. 1974.
- KANAL, L.N.; CHANDRASEKARAN, B. On dimensionality and sample size in statistical pattern classification. *Pattern Recognition*, 3:225-234, Oct. 1971.
- KAUTH, R.J.; THOMAS, G.S. The tasselled cap-A graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: SYMPOSIUM ON MACHINE PROCESSING OF REMOTELY SENSED DATA, West Lafayette, IN, 1976. *Proceedings*. Purdue University, 1976.
- KETTIG, R.L.; LANDGREBE, D.A. Classification of multispectral image data by extraction and classification of homogeneous objects. *IEEE Transactions on geoscience electronics*, GE-4(1), Jan. 1976.
- LANDGREBE, D.A. Analysis technology for land remote sensing. *Proceedings of the IEEE*, 69(5):628-642, Mar. 1981.

- LINDE, Y.; BUZO, A.; GRAY, R.M. An algorithm for vector quantizer design. *IEEE Transactions on communications*, COM-28(1):84-95, Jan. 1980.
- MAC QUEEN, J. Some methods for classification and analysis of multivariate observations. In: SYMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, 5., Los Angeles, CA, University of California. *Proceedings*, p. 281-297, 1967.
- MATUSITA, K. Decision rules, based on the distance for problems of fit, two samples and estimation. *Annals of mathematical statistics*, 26:631-640, 1955.
- MISRA, P.N.; WHEELER, S.G. Crop classification with Landsat multispectral scanner data, *Pattern Recognition*, 10(1):1-13, Jan. 1978.
- MOIK, J.G. *Digital processing of remotely sensed images*. Washington, DC, US Government Printing Office, 1980.
- ROSENBLAT, F. *Principles of neurodynamics: perceptrons and the theory of brain mechanisms*, Washington, DC, Spartan Books, 1961.
- ROSENFELD, A.; HUMMEL, R.; ZUCKER, S. Scene labeling by relaxation algorithms. *IEEE Transactions on systems, man and cybernetics*, SMC-6(6):420-433, June, 1976.
- SHLIEN, S.; SMITH, A. *Practical aspects related to automatic classification of Landsat imagery using look-up tables*. Canadian Center for Remote Sensing, Feb. 1975. (Research Report, 75-2).
- SWAIN, P.H. Bayesian classification in a time-varying environment, *IEEE transactions on systems, man and cybernetics*, SMC-8(12):879-888, Dec. 1978.
- SWAIN, P.H.; DAVIS, S.M. ed. *Remote sensing - the quantitative approach*, New York, NY, Mc Graw Hill, 1978.
- SWAIN, P.H.; SIEGEL, H.J.; SMITH, B.W. Contextual classification of multispectral remote sensing data using a multispectral system. *IEEE Transactions on Geoscience Electronics*, GE-18(2):197-203, 1980.
- TOU, J.T.; GONZALEZ, R.C. *Pattern recognition principles*. Reading, Mass., Addison Wesley, 1974.

- TOU, J.T.; HEYDORN, R.P. Some approaches to optimum feature extraction.
In: Tou, J.T., ed. *Computer and information sciences-II*, New York,
NY, Academic, v.2, 1967.
- TOUSSAINT, G.T. Bibliography on estimation of misclassification. *IEEE
Transactions on Information Theory*, IT-20(4):472-479, July, 1974.
- VELASCO, F.R.D. Thresholding using the ISODATA clustering Algorithm,
IEEE Transactions on systems, man and cybernetics, SMC-10, 771-774,
Nov. 1980.
- YOUNG, T.Y.; CALVERT, T.W. *Classification, estimation and pattern
recognition*. New York, NY, American Elsevier, 1974.

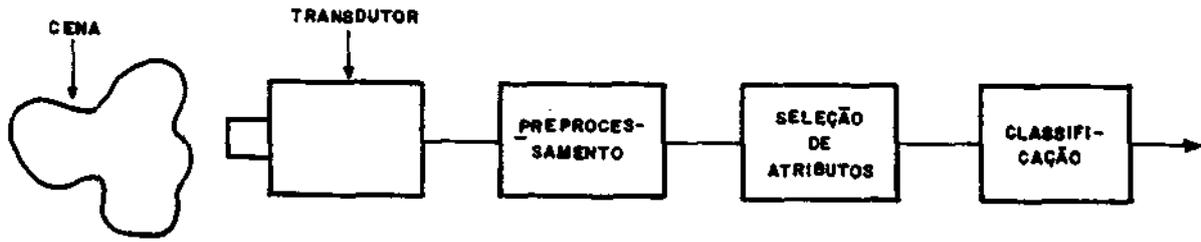


Fig. 10.1 - Esquema geral de um processo de classificação de imagens

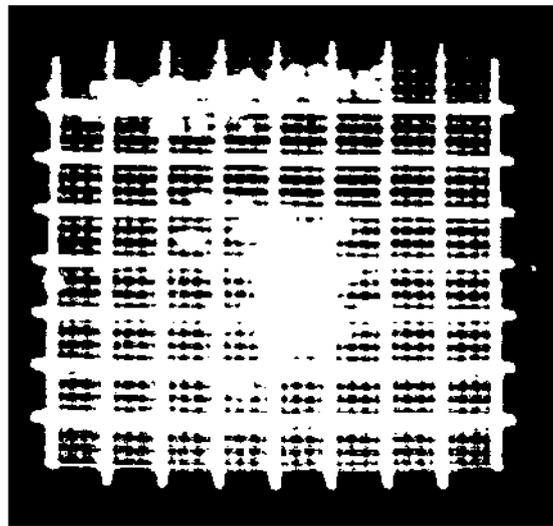


Fig. 10.2 - Distribuição das amostras de duas classes em dois canais de uma imagem LANDSAT.

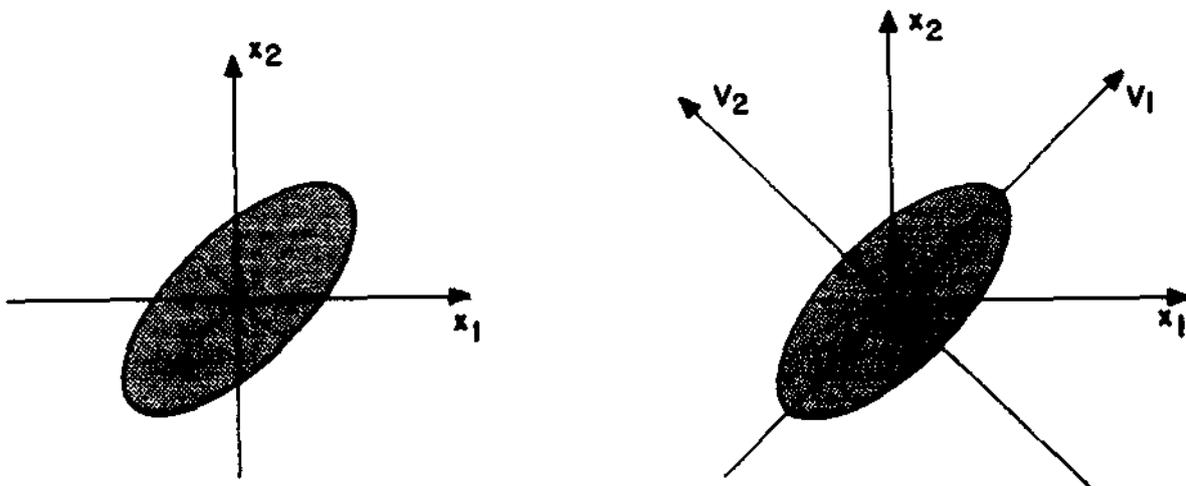


Fig. 10.3 - Interpretação geométrica da Transformação de Karhunen-Loeve em duas dimensões.

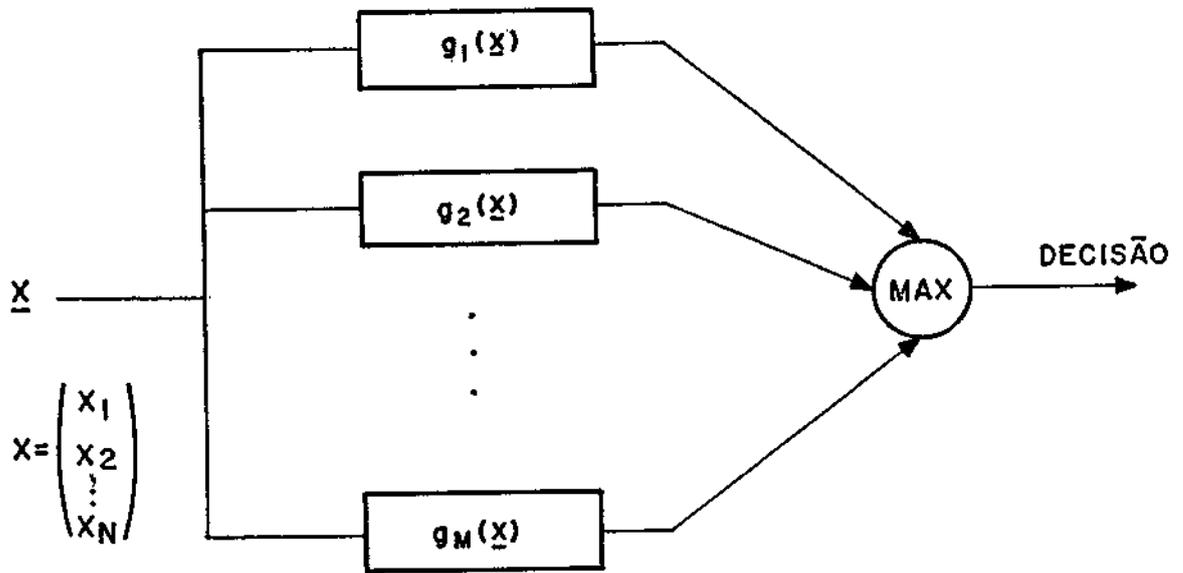


Fig. 10.4 - Decisão por funções discriminantes.

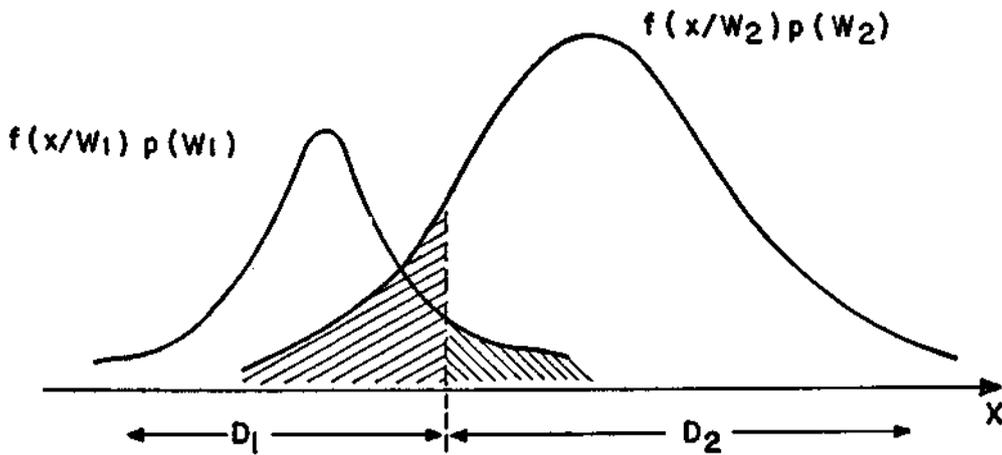


Fig. 10.5 - Classificação estatística: cálculo das probabilidades de erro no caso gaussiano unidimensional.

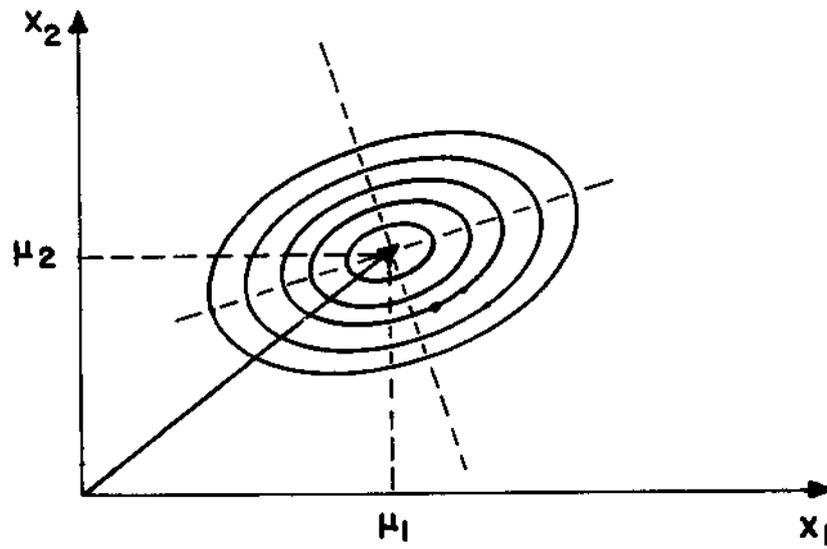


Fig. 10.6 - Curvas de nível da função densidade de probabilidade gaussiana.

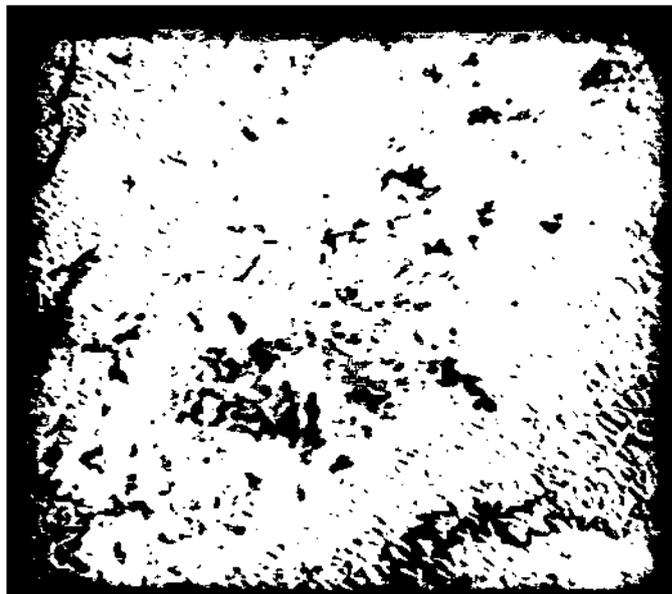


Fig. 10.7 - Imagem LANDSAT - região de São José dos Campos, SP.

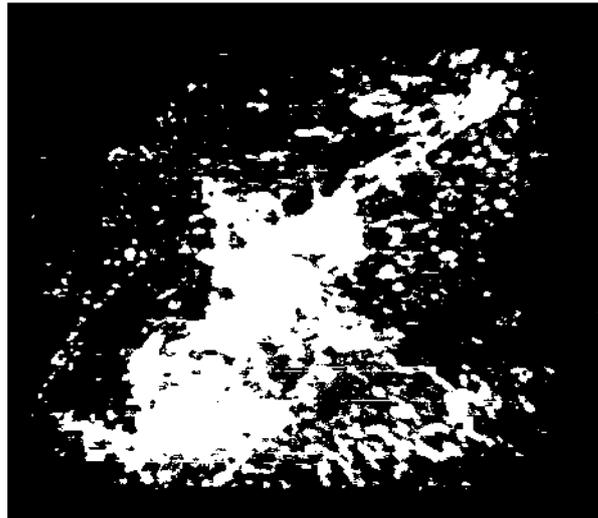


Fig. 10.8 - Resultado da classificação da Figura 10.7 pelo algoritmo de máxima verossimilhança com a hipótese gaussiana: classe "área urbana"

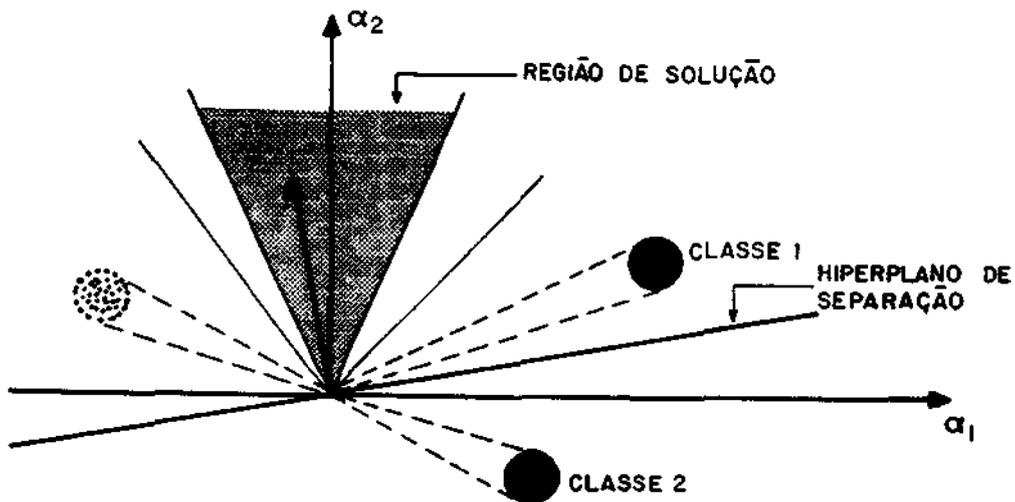


Fig. 10.9 - Hiperplano de separação entre duas classes.

APÊNDICE

A área de processamento de imagens deriva suas bases das teorias de sistemas lineares, de análise de Fourier, de transformações discretas em imagens e de campos aleatórios. Neste Apêndice serão apresentadas as principais noções dessas áreas, necessárias para a compreensão das deduções apresentadas no texto. Para maior detalhamento, o leitor deverá consultar a bibliografia apresentada no final deste Apêndice a qual serviu de base para a sua elaboração.

A.1 - Sistemas Lineares

Diz-se que uma operação O , que transforma uma imagem em outra, é linear se

$$O [af + bg] = a O[f] + b O[g] \quad (A.1)$$

para todos os pares de imagens f e g e todas as constantes a e b .

O conceito de fonte pontual é útil e pode ser expresso pela função de Dirac $\delta(x,y)$, definida por:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \delta(x,y) dx dy = 1, \quad (A.2)$$

tal que $\delta(x,y) = 0$ para todo $(x,y) \neq (0,0)$, e $\delta(0,0) \rightarrow \infty$.

A chamada propriedade de esquadramento da função $\delta(x,y)$ é definida por:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) \delta(x-\xi, y-\eta) d\xi d\eta = f(x,y). \quad (A.3)$$

A resposta de um sistema linear a uma fonte pontual no plano é chamada Função de Espalhamento Pontual (F.E.P.) e é denotada por $h(x,y)$.

Uma operação linear é denominada invariante no espaço se a resposta a uma fonte pontual localizada em (ξ,η) isto é, $\delta(x-\xi,y-\eta)$, é dada por $h(x-\xi,y-\eta)$.

Usando a Equação A.3 e utilizando a hipótese de linearidade da operação O , obtêm-se:

$$\begin{aligned} O[f(x,y)] &= O \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) \delta(x-\xi,y-\eta) d\xi d\eta \right], \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) O[\delta(x-\xi,y-\eta)] d\xi d\eta, \end{aligned} \quad (A.4)$$

que resulta em:

$$O[f(x,y)] = g(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) h(x-\xi,y-\eta) d\xi d\eta. \quad (A.5)$$

A Equação A.5 expressa a chamada convolução das funções f e g , denotada às vezes por $f * g$. A operação de convolução tem importância em problemas de realce, restauração e reconstrução de imagens.

Uma operação análoga à de convolução é a de correlação que pode ser definida entre duas funções f e g (correlação cruzada):

$$R_{fg}(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) g(x+\xi, y+\eta) d\xi d\eta, \quad (A.6)$$

ou entre uma função f e ela mesma (autocorrelação):

$$R_{ff}(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) f(x+\xi, y+\eta) d\xi d\eta. \quad (A.7)$$

Uma aplicação da correlação cruzada reside no problema de registro translacional de imagens.

A.2 - Análise de Fourier

A Transformada Bidimensional de Fourier de uma função $f(x,y)$ pode ser definida por:

$$F(u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) e^{-j2\pi(ux+vy)} dx dy. \quad (A.7)$$

A Transformada Inversa é obtida por:

$$f(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u,v) e^{+j2\pi(ux+vy)} du dv. \quad (A.8)$$

De modo geral, a função $F(u,v)$ assume valores complexos.

Algumas propriedades da Transformada Bidimensional de Fourier, que são úteis em processamento de imagens, são apresentadas a seguir:

a) linearidade:

$$F [a f(x,y) + b g(x,y)] = a F(u,v) + b G(u,v); \quad (A.9)$$

b) escalonamento:

$$F [f(\xi x + \eta y)] = \frac{1}{|\xi\eta|} F \left[\frac{u}{\xi}, \frac{v}{\eta} \right]; \quad (A.10)$$

c) deslocamento no domínio do espaço:

$$F [f(x-\xi, y-\eta)] = F(u,v) e^{-j2\pi(u\xi + v\eta)}; \quad (A.11)$$

d) deslocamento no domínio da frequência;

$$F \left[f(x,y) e^{j2\pi(u_0x + v_0y)} \right] = F(u - u_0, v - v_0); \quad (A.12)$$

e) rotação de 180°

$$F [F^{-1} [f(x,y)]] = f(-x, -y); \quad (A.13)$$

f) convolução no domínio do espaço:

$$F \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) g(x-\xi, y-\eta) d\xi d\eta \right] = F(u,v) \cdot G(u,v); \quad (A.14)$$

g) convolução no domínio da frequência:

$$F [f(x,y) \cdot g(x,y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(s,t) G(u-s, v-t) ds dt; \quad (A.15)$$

h) correlação cruzada:

$$F \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) g(x+\xi, y+\eta) d\xi d\eta \right] = F^*(u,v) G(u,v); \quad (A.16)$$

i) autocorrelação:

$$F \left[\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\xi,\eta) f(x+\xi, y+\eta) d\xi d\eta \right] = |F(u,v)|^2; \quad (A.17)$$

j) teorema de Parseval:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) g^*(x,y) dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u,v) \cdot G^*(u,v) du dv. \quad (A.18)$$

No caso particular em que $f(x,y) = g(x,y)$ tem-se:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x,y)|^2 dx dy = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |F(u,v)|^2 du dv. \quad (A.19)$$

Se a função $f(x,y)$ tem simetria radial, ou seja, se $f(x,y) = f[(x^2 + y^2)^{1/2}] = f(r)$, a Transformada Bidimensional de Fourier pode ser dada por:

$$F(u,v) = F(w) = 2\pi \int_0^{\infty} r J_0(2\pi r w) f(r) dr, \quad (A.20)$$

e a Transformada Inversa \bar{e} obtida por:

$$f(r) = 2\pi \int_0^{\infty} w J_0(2\pi r w) F(w) dw, \quad (A.21)$$

onde $w = (u^2 + v^2)^{1/2}$ \bar{e} a frequência radial espacial e $J_0(r)$ \bar{e} a função de Bessel do 1º tipo de ordem zero. As Equações A.20 e A.21 representam a chamada Transformada de Hankel e sua inversa e têm grande importância em processamento ótico de imagens.

A.3 - Transformações Discretas em Imagens

Uma imagem pode ser representada por uma matriz de $M \times N$ números. Uma transformação discreta linear unitária numa imagem pode ser expressa pela Equação:

$$F(m,n) = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} \vartheta_{mn}(j,k) f(j,k) \quad (\text{A.22})$$

$$m = 0, 1, \dots, M-1$$

$$n = 0, 1, \dots, N-1$$

e a Transformada Inversa \bar{e} obtida por:

$$f(j,k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \vartheta_{mn}^*(j,k) F(m,n) \quad (\text{A.23})$$

$$j = 0, 1, \dots, M-1$$

$$k = 0, 1, \dots, N-1$$

Os elementos $f(j,k)$ e $F(m,n)$ constituem as matrizes $(M \times N)$ $\underline{[f]}$ e $\underline{[F]}$, respectivamente e $\vartheta_{mn}(j,k)$ (ou seu complexo conjugado $\vartheta_{mn}^*(j,k)$) \bar{e} um elemento de um operador quadridimensional.

Se for feita uma ordena \tilde{c} o lexicogr \tilde{a} fica dos MN componentes das matrizes $\underline{[f]}$ e $\underline{[F]}$, as Equa \tilde{c} o \tilde{e} s A.22 e A.23 podem ser expressas por:

$$\underline{F} = \underline{\vartheta} \underline{f}, \quad (\text{A.24})$$

$$\underline{f} = \underline{\vartheta}^{*T} \underline{F}, \quad (\text{A.25})$$

onde $\underline{\vartheta}$ \bar{e} uma matriz $(MN \times MN)$ ($\underline{\vartheta}^{*T}$ \bar{e} sua conjugada transposta) e \underline{F} e \underline{f} s \tilde{a} o vetores $(MN \times 1)$.

A matriz de transforma \tilde{c} o \bar{e} considerada separ \tilde{a} vel se para ca da (m,n) a seguinte fatoriza \tilde{c} o for poss \tilde{i} vel:

$$\vartheta_{mn}(j,k) = p_m(j) q_n(k). \quad (\text{A.26})$$

Uma transforma \tilde{c} o bidimensional separ \tilde{a} vel pode ser computada em duas etapas, envolvendo cada uma delas uma transforma \tilde{c} o unidimensional. Inicialmente as linhas de $f(i,j)$ s \tilde{a} o transformadas por:

$$F(m,k) = \sum_{j=0}^{M-1} f(j,k) P_m(j) \quad (A.27)$$

e, em seguida, as colunas de $F(m,k)$ são transformadas por:

$$F(m,n) = \sum_{k=0}^{N-1} F(m,k) q_n(k). \quad (A.28)$$

Definindo a matriz \underline{P} ($M \times M$) por:

$$P(m,j) = P_m(j) \quad (A.29)$$

e a matriz \underline{Q} ($N \times N$) por:

$$Q(n,k) = q_n(k),$$

pode-se expressar a Equação A.22 por:

$$[\underline{F}] = \underline{P}^T [\underline{f}] \underline{Q} \quad (A.30)$$

e a Transformada Inversa dada por A.23 pode ser dada por:

$$[\underline{f}] = \underline{P}^* [\underline{F}] \underline{Q}^{*T}. \quad (A.31)$$

Várias transformações separáveis unitárias têm sido empregadas em processamento digital de imagens, particularmente nas áreas de restauração, reconstrução e codificação. Entre elas podem ser mencionadas as Transformadas Discretas de Fourier, do cosseno, e a de Hadamard, que possuem implementações por meio de algoritmos rápidos. A Transformada de Karhunen-Loeve é, em geral, uma transformação não-separável que depende das características estatísticas da imagem e que desempenha importante papel em codificação e em classificação de imagens.

A.3.1 - Transformada Discreta de Fourier

Na Transformada Discreta de Fourier (em inglês Discrete Fourier Transform-DFT) as matrizes unitárias de transformação são dadas por:

$$p_m(j) = \frac{1}{M} e^{-i2\pi j \frac{m}{M}} \quad j, m = 0, 1, \dots, M-1 \quad (\text{A.32})$$

e

$$q_n(k) = \frac{1}{N} e^{-i2\pi k \frac{n}{N}} \quad k, n = 0, 1, \dots, N-1, \quad (\text{A.33})$$

onde $i = \sqrt{-1}$. Assim sendo, a transformada direta tem a forma:

$$F(m, n) = \frac{1}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f(j, k) e^{-i2\pi \left(j \frac{m}{M} + k \frac{n}{N} \right)}$$

$$m = 0, 1, \dots, M-1$$

$$n = 0, 1, \dots, N-1. \quad (\text{A.34})$$

A transformada inversa \tilde{e} dada por:

$$f(j, k) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(m, n) e^{i2\pi \left(j \frac{m}{M} + k \frac{n}{N} \right)}$$

$$j = 0, 1, \dots, M-1$$

$$k = 0, 1, \dots, N-1. \quad (\text{A.35})$$

As propriedades (a) - (j) listadas neste Apêndice para a Transformada de Fourier no caso contínuo também são válidas para o caso discreto.

Além disso, as propriedades de periodicidade do operador exponencial implicam:

$$F(m, -n) = F(m, N-n), \quad (\text{A.36})$$

$$F(-m, n) = F(M-m, n), \quad (\text{A.37})$$

$$F(-m, -n) = F(M-m, N-n), \quad (\text{A.38})$$

bem como:

$$f(j, -k) = f(j, N-k), \quad (A-39)$$

$$f(-j, k) = f(M-j, k), \quad (A-40)$$

$$f(-j, -k) = f(M-j, N-k). \quad (A-41)$$

Esses resultados implicam que, estendendo-se as matrizes $[f(j,k)]$ e $[F(m,n)]$ além do domínio original dado por $[0 \leq j, m \leq M - 1)$ e $[0 \leq k, n \leq N-1]$, obtêm-se repetições periódicas dessas matrizes.

O fato de as imagens serem representadas por valores reais implica a propriedade de simetria conjugada, ou seja:

$$F(m,n) = F^* (-m, -n). \quad (A.42)$$

As propriedades expressas pelas Equação A.36, A.37, A.38 e A.42, quando combinadas, implicam que:

$$|F(m,n)| = |F(M-m, N-n)|, \quad (A.43)$$

$$|F(m, N-n)| = |F(M-m,n)|. \quad (A.44)$$

Assim sendo ao se fazer uma exibição do módulo da Transformada Discreta de Fourier Bidimensional, as frequências correspondentes aparecerão em cantos opostos, dificultando a interpretação. Todavia, se for feita antes a multiplicação pelo fator $e^{i\pi(j+k)} = (-1)^{(j+k)}$, pela propriedade de translação, a frequência (0,0) da transformada será deslocada para o centro da exibição, isto é, $(M/2, N/2)$, facilitando assim a interpretação do resultado.

Devido à separabilidade da DFT, sua computação no caso bidimensional envolve a transformação das linhas e em seguida a transformação das colunas, o que normalmente exige um passo adicional de transposição da imagem, que usualmente é armazenada em memória secundária.

Cada transformada unidimensional é executada pelo algoritmo da Transformada Rápida de Fourier (em inglês "Fast Fourier Transform" - FFT), que reduz o número de operações de N^2 para $N \log N$, o que se torna substancial para grandes valores de N (para uma imagem com 512 elementos numa direção essa redução é aproximadamente de um fator de 57).

A.3.2 - Transformada Discreta do Cosseno

A Transformada Discreta do Cosseno Bidimensional é definida por:

$$F(0,0) = \frac{2}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f(j,k), \quad (A.45)$$

$$F(m,n) = \frac{4}{MN} \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} f(j,k) \cdot \cos\left\{\frac{2j+1}{2M} \Pi m\right\} \cdot \cos\left\{\frac{2k+1}{2N} \Pi n\right\} \quad (A.46)$$

$$m = 1, 2, \dots, M-1$$

$$n = 1, 2, \dots, N-1.$$

A transformada inversa é obtida por:

$$f(j,k) = \frac{1}{2} F(0,0) + \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} F(m,n) \cos\left\{\frac{2j+1}{2M} \Pi m\right\} \cdot \cos\left\{\frac{2k+1}{2N} \Pi n\right\} \quad (A.47)$$

$$j = 0, 1, \dots, M-1$$

$$k = 0, 1, \dots, N-1.$$

Como tanto a transformação direta como a inversa são separáveis, elas podem ser implementadas por transformações unidimensionais. Pode-se mostrar também que a Transformada Discreta do Cosseno pode ser computada pelo algoritmo de FFT. A Transformada do Cosseno tem sido recentemente objeto de grande interesse na área de codificação de imagens.

A.3.3 - Transformada Discreta de Hadamard

A Transformada Discreta de Hadamard é obtida quando as matrizes P e Q na Equação A.30 são as chamadas matrizes de Hadamard. Estas matrizes são simétricas, formadas por elementos +1 e -1 e colunas e linhas ortogonais entre si.

Existem várias formulações possíveis para a Transformada Discreta de Hadamard Bidimensional. Uma delas, definida para uma imagem de $N \times N$ elementos, onde $N = 2^n$, é dada por:

$$F(m,n) = \frac{1}{N} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j,k) (-1)^{\sum_{i=0}^{n-1} [b_i(j)b_i(m) + b_i(k)b_i(n)]} \quad (A.48)$$

e a transformada inversa é obtida por uma expressão idêntica à anterior, trocando os papéis de $F(m,n)$ e $f(j,k)$. Na Equação A.48 a somatória no expoente de (-1) é feita em aritmética módulo 2 e $b_i(z)$ denota o $i^{\text{ésimo}}$ bit na representação binária de z .

A principal utilização da Transformada Discreta de Hadamard Bidimensional tem sido em codificação de imagens.

A.3.4 - Transformada Discreta de Karhunen-Loeve

A Transformação de Karhunen-Loeve de uma imagem de $N \times N$ elementos tem a forma geral das transformações discretas lineares unitárias, ou seja:

$$F(m,n) = \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} \phi_{mn}(j,k) f(j,k), \quad (A.49)$$

onde o cerne $\phi_{mn}(j,k)$ satisfaz a Equação:

$$\lambda(m,n) \cdot \phi_{mn}(j,k) = \sum_{j'=0}^{N-1} \sum_{k'=0}^{N-1} C_f(j,k; j',k') \cdot \phi_{mn}(j',k'), \quad (A.50)$$

onde $C_f(k, j', k')$ denota a função de covariância da matriz que descreve a imagem e $\lambda(m,n)$, é constante para um valor fixo de (m,n) . O conjunto de funções que definem o cerne são as funções próprias da função de covariância e $\lambda(m,n)$ representam os valores próprios dessa mesma função.

Se a função de covariância é separável, isto é, se:

$$C_f(j,k, j', k') = C_c(j,j') \cdot C_l(k,k'), \quad (A.51)$$

então o cerne da Transformação de Karhunen-Loeve também é separável e tem-se:

$$\phi_{mn}(j,k) = \phi_c(j,m) \cdot \phi_l(k,n) \quad (A.52)$$

e são válidas as seguintes equações do tipo valor próprio-função própria para as linhas e colunas:

$$\lambda_l(n) \phi_l(k,n) = \sum_{k'=0}^{N-1} C_l(k,k') \phi_l(k',n), \quad (A.53)$$

$$\lambda_c(m) \phi_c(j,m) = \sum_{j'=0}^{N-1} C_c(j,j') \phi_c(j',m). \quad (A.54)$$

Se a imagem e sua transformada são expressas por vetores, usando uma notação lexicográfica, a Transformação de Karhunen-Loeve e sua inversa têm a forma:

$$\underline{F} = \underline{\phi} \underline{f}, \quad (A.55)$$

$$\underline{f} = \underline{\phi}^T \underline{F}. \quad (A.56)$$

A matriz de transformação $\underline{\phi}$ satisfaz a Equação:

$$\underline{\phi} \underline{C}_f = \underline{\Lambda} \underline{\phi}, \quad (A.57)$$

onde \underline{C}_f é a matriz de covariância do vetor \underline{f} e $\underline{\Lambda}$ é uma matriz diagonal onde os N^2 elementos não nulos são os valores próprios da matriz \underline{C}_f .

A Transformação de Karhunen-Loeve tem sido utilizada no domínio espacial da imagem em codificação mas também no caso de imagens ou multitemporais de multitemporais, principalmente para extração de atributos em classificação de imagens, visando a redução de dimensionalidade. Nesse caso o vetor \underline{f} tem a dimensão do número de canais e a transformação é aplicada em cada elemento vetorial (multiespectral ou multitemporal, por exemplo) da imagem.

Do ponto de vista de redução de dimensionalidade, a Transformada de Karhunen-Loeve possui propriedades de otimalidade. Assim, se o vetor \underline{f} na Equação A.56 possui dimensionalidade P , essa equação pode ser expressa por:

$$\underline{f} = \sum_{n=1}^P F_n \underline{\phi}_n, \quad (\text{A.58})$$

onde $\underline{\phi}_n$ são os vetores-coluna de $\underline{\phi}^T$ (ou linhas de $\underline{\phi}$).

A Equação A.55, por sua vez pode ser expressa por:

$$F_n = \underline{\phi}_n^T \underline{f} \quad n = 1, 2, \dots, P. \quad (\text{A.59})$$

Em codificação ou extração de atributos, \underline{f} deve ser aproximado no sentido de erro médio quadrático (ϵ) com o menor número de coeficientes F_n , $n = 1, 2, \dots, N < P$ possível.

O erro médio quadrático pode ser dado por:

$$\epsilon = E \left[\left\| \underline{f} - \sum_{n=1}^N F_n \underline{\phi}_n \right\|^2 \right] \quad (\text{A.60})$$

e deve ser minimizado pela escolha dos vetores $\underline{\phi}_n$, $n = 1, 2, \dots, N$.

Utilizando a Equação A.59 e a propriedade de ortonormalidade dos vetores $\{\underline{\phi}_n\}$ (que advém de fato de a matriz $\underline{\phi}$ ser unitária), tem-se:

$$\underline{\epsilon} = \sum_{n=N+1}^P \underline{\phi}_n^T \underline{R}_f \underline{\phi}_n, \quad (\text{A.61})$$

onde \underline{R}_f é a matriz de correlação do vetor \underline{f} , relacionado à matriz de covariância \underline{C}_f por:

$$\underline{C}_f = \underline{R}_f - \underline{m}_f \underline{m}_f^T \quad (\text{A.62})$$

onde \underline{m}_f é o vetor média do vetor aleatório \underline{f} . Sob a restrição de norma unitária dos vetores $\underline{\phi}_n$, pode-se obter o mínimo da Expressão A.61. pela utilização do método dos multiplicadores de Lagrange, obtendo-se:

$$\underline{R} \underline{\phi}_n = \lambda_n \underline{\phi}_n, \quad (\text{A.63})$$

que é uma equação do tipo vetor próprio - valor próprio da matriz de correlação. (A Equação A.63 corresponde à Equação A.50 se se supuzer que o vetor \underline{f} tem média nula; nesse caso a matriz de covariância e de correlação coincidem. Caso isso não aconteça, pode-se definir uma transformação envolvendo um vetor de média nula por $\underline{F} = \underline{\phi}(\underline{f} - \underline{m}_f)$).

O erro médio quadrático é dado por:

$$E = \sum_{n=N+1}^P \lambda_n, \quad (\text{A.64})$$

ou seja, ele será minimizado se forem eliminadas as componentes correspondentes aos menores valores próprios (que são todos reais e positivos, supondo \underline{R} uma matriz positiva definida).

Além disso, pode-se mostrar que a matriz de correlação de \underline{F} é dada por:

$$R_F = E [\underline{F} \underline{F}^T] = \underline{\Lambda}, \quad (A.65)$$

onde $\underline{\Lambda}$ é a matriz diagonal que foi definida na Equação A.57, ou seja, as componentes do vetor transformado \underline{F} são não-correlacionadas e suas variâncias são os valores próprios envolvidos ($\lambda_n, n=1,2,\dots,P$).

Como a transformada é unitária, pode-se ainda mostrar que:

$$\sum_{n=1}^P (\sigma_n^f)^2 = \sum_{n=1}^P \lambda_n, \quad (A.66)$$

onde σ_n^f são as variâncias das componentes do vetor \underline{f} .

A.4 - Noções de Campos Aleatórios

Em várias aplicações, uma imagem deve ser representada como uma realização de um campo aleatório, ao invés de uma função determinística de duas variáveis. Campos aleatórios podem representar estatisticamente classes de imagens, como imagens multiespectrais ou multitemporais. A aleatoriedade pode advir da variação própria da cena ou das flutuações estatísticas dos sensores.

Um campo aleatório bidimensional pode ser representado por $f(x,y,w_i)$ onde w_i é um evento no espaço de amostragem Ω . Para um dado valor de (x,y) , $f(x,y,w_i)$ é uma variável aleatória. Para um dado valor de w_i , $f(x,y,w_i)$ é uma função no plano x,y . Fixados x,y e w_i , $f(x,y,w_i)$ é um número real. Por simplificação, o campo aleatório $f(x,y,w_i)$ que representa uma classe de imagens será representado daqui por diante por $f(x,y)$.

Um campo aleatório é completamente determinado pela função densidade de probabilidade $p_f(z_1 \dots z_n; x_1, y_1, \dots, x_n, y_n)$. Todavia, densidades de alta ordem são difíceis de obter. Algumas vezes a densidade de 1ª ordem $p_f(z; x,y)$ tem sido usado para caracterizar imagens (como as densidades gaussianas, exponenciais ou de Rayleigh, etc).

Se os momentos até 2ª ordem do campo aleatório, são independentes do espaço, diz-se que campo é estacionário no sentido amplo. Nesse caso o valor esperado, definido por:

$$\mu_f = E [f(x,y)] = \int_{-\infty}^{+\infty} z p_f(z; x,y) dz. \quad (A.67)$$

é constante e a autocorrelação só depende da separação entre os dois elementos de imagem, isto é:

$$E [f(x + \xi, y + n) f(x,y)] = R_{ff}(\xi, n). \quad (A.68)$$

A função de correlação pode ser também invariante a rotações. Nesse caso:

$$R_{ff}(\xi, n) = R_{ff} [(\xi^2 + n^2)^{1/2}] \quad (A.69)$$

e o campo aleatório é chamado isotrópico.

Uma outra forma frequentemente utilizada para representar a função de correlação de campos aleatórios é do tipo:

$$R_{ff}(\xi, n) = (R_{ff}(0,0) - \mu_f^2) e^{-\alpha|\xi| - \beta|n|} + \mu_f^2. \quad (A.70)$$

Se $\mu_f = 0$ essa função de correlação será separável nas direções horizontal e vertical.

Dados dois campos aleatórios f e g , diz-se que eles são conjuntamente estacionários no sentido amplo se sua correlação cruzada é função apenas da separação dos elementos, isto é, se:

$$E [f(x + \xi, y+n) g(x,y)] = R_{fg}(\xi,n). \quad (A.71)$$

A função de covariância de f e g é expressa por:

$$C_{fg}(\xi, n) = R_{fg}(\xi, n) - \mu_f \mu_g. \quad (A.72)$$

Os campos serão considerados não-correlacionados se:

$$c_{fg}(\xi, n) = 0 \quad (A.73)$$

e ortogonais se:

$$R_{fg}(\xi, n) = 0. \quad (A.74)$$

Pode-se definir a função densidade espectral de potência de um campo aleatório estacionário no sentido amplo como a Transformada de Fourier Bidimensional de sua função de correlação, isto é:

$$S_{ff}(u, v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} R_{ff}(\xi, n) e^{-j2\pi(\xi u + n v)} d\xi dn. \quad (A.75)$$

Pode-se mostrar que, passando-se um campo aleatório $f(x, y)$ estacionário no sentido amplo por um filtro linear invariante no espaço com função de transferência $H(u, v)$, o campo aleatório $g(x, y)$ também será estacionário no sentido amplo e sua função densidade espectral de potência será dada por:

$$S_{gg}(u, v) = |H(u, v)|^2 S_{ff}(u, v). \quad (A.76)$$

As definições de μ_f e $R_{ff}(\xi, n)$ para um campo aleatório estacionário no sentido amplo envolvem médias sobre todas as possíveis realizações desse processo (o chamado "ensemble"). Tais médias são dificilmente obtidas na prática. Sob determinadas condições é possível usar médias espaciais de uma realização como estimativas satisfatórias das médias mencionadas anteriormente. Nessas condições diz-se que o campo aleatório é ergódico.

REFERÊNCIAS BIBLIOGRÁFICAS

- GONZALEZ, R.C.; WINTZ, P. *Digital image processing*. Reading, MA., Addison Wesley, 1977.
- MOIK, J.G. *Digital Processing of Remotely Sensed Images*. Washington, DC, U.S. Government Printing Office, 1980.
- PRATT, W.K. *Digital image processing*. New York, NY, John Wiley, 1978.
- ROSENFELD, A.; KAK, A.C. *Digital picture processing*. 2ª Ed., New York, NY, Academic, 1982.