

INPE-7502-TDI/717

## REGISTRO AUTOMÁTICO DE IMAGENS DE SENSORIAMENTO REMOTO BASEADO EM MÚLTIPLAS RESOLUÇÕES

Leila Maria Garcia Fonseca

Tese de Doutorado em Computação Aplicada, orientada pelo Dr. Max Henrique Machado Costa e pelo Dr. Gerald Jean Francis Banon, aprovada em 26 de fevereiro de 1999.

INPE São José dos Campos 2000

### Publicado por:

Coordenação de Ensino, Documentação e Programas Especiais - CEP

Instituto Nacional de Pesquisas Espaciais - INPE Caixa Postal 515

12201-970 - São José dos Campos - SP - Brasil

Fone: (012) 345.6911 Fax: (012) 345.6919

E-Mail: marciana@sid.inpe.br

- Solicita-se intercâmbio
- Exchange welcome
- Si sollecita intercambio
- Echange souhaité
- Mann bittet un Austausch
- Solicitamos intercambio
- -Просим овмену
- 数进并介点按
- -出股物交換《不解》

Publicação Externa - É permitida sua reprodução para interessados.

## MINISTÉRIO DA CIÊNCIA E TECNOLOGIA INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-7502-TDI/717

# REGISTRO AUTOMÁTICO DE IMAGENS DE SENSORIAMENTO REMOTO BASEADO EM MÚLTIPLAS RESOLUÇÕES

Leila Maria Garcia Fonseca

Tese de Doutorado em Computação Aplicada, orientada pelo Dr. Max Henrique Machado Costa e pelo Dr. Gerald Jean Francis Banon, aprovada em 26 de fevereiro de 1999.

INPE São José dos Campos 2000 621.376.5 : 528.711.7

FONSECA, L. M. G.

Registro automático de imagens de sensoriamento remoto baseado em múltiplas resoluções / L.M.G.Fonseca – São José dos Campos: INPE, 1999.

219p. – (INPE-7502-TDI/717).

1.Registro de imagens. 2.Transformada wavelet. 3.Resolução. 4.Sensoriamento remoto. 5.Satélites artificiais. 6.Correlação. I.Título.

Aprovada pela Banca Examinadora em cumprimento a requisito exigido para a obtenção do Título de **Doutora** em **Computação Aplicada.** 

Dr. Luciano Vieira Dutra	Presidente
Dr. Gerald Jean Francis Banon	Orientador
Dr. Max Henrique Machado Costa	Orientador
Dr. Gilberto Câmara Neto	Membro da Banca
Dr. Nelson Delfino A. Mascarenhas	Membro da Banca Convidado
Dr. Fernando Sakane	Membro da Banca Convidado

Candidato (a): Leila Maria Garcia Fonseca

Ao meu querido esposo André e as minhas adoráveis filhas Laura e Julia, pelo apoio, compreensão, incentivo e o amor incondicional.

#### **AGRADECIMENTOS**

Primeiramente, agradeço aos meus orientadores Dr. Max Costa e Dr. Banon, pelo encorajamento, apoio e, principalmente, pela serenidade e confiança com que orientaram este trabalho.

Minha sincera gratidão ao Dr. B. S. Manjunath, com quem iniciei as pesquisas deste trabalho na Universidade de Santa Barbara, California (UCSB), nos períodos de 03/94 a 07/95 e 05/97 a 07/97. Estendo os meus agradecimentos aos colegas na UCSB, Charlies Kenney, Wei, Eduardo, Markus e Norbert pela amizade e apoio.

Eu gostaria de agradecer aos membros da Banca Examinadora pelas valiosas sugestões que melhoraram significativamente este trabalho. Agradecimentos especiais ao Dr. Mascarenhas, orientador de minha dissertação de Mestrado, por ter me dado a honra de tê-lo como membro da Banca, mesmo estando de férias.

Aos meus colegas de trabalho da DPI, obrigada pela amizade e o excelente ambiente de trabalho. Em especial, agradeço a Sueli pelo carinho e amizade sincera. Agradecimentos ao Sidnei, Camilo, Julio D'Alge e Guaraci pelas contribuições e valiosas trocas de idéias. Obrigada ao Dr. Gilberto Câmara e Ubirajara Freitas que, como chefes da divisão, me permitiram realizar o programa de doutorado, disponibilizando material e equipamentos. Agradecimento ao Diehl pelo suporte computacional, ao Dr. Diógenes S. Alves e ao Sergio Faria por disponibilizarem algumas imagens usadas neste trabalho. Especial gratidão ao Sergio Faria pela amizade e incentivo.

Este trabalho não teria sido possível sem o apoio, carinho, compreensão e paciência do meu esposo André e minhas filhas Laura e Julia. Aproveito esta oprotunidade para expressar minha profunda gratidão a eles.

Finalmente, gostaria de agradecer ao CNPq pelo suporte financeiro durante as pesquisas realizadas na UCSB.

#### **RESUMO**

Registro de imagens é uma das operações básicas de processamento de imagens na área de sensoriamento remoto. Com o aumento do número de imagens coletadas todos os dias por diferentes sensores, o registro automático de imagens multi-sensores, multi-espectrais e multi-temporais tem se tornado um assunto muito importante. Nesta tese é proposto um eficiente algoritmo para o registro de imagens multi-temporais com respostas espectrais similares. O procedimento é completamente automático e baseia-se nos valores de níveis de cinza das imagens e nos máximos locais da transformada wavelet. O algoritmo é realizado nas resoluções progressivamente mais altas, permitindo assim uma maior eficiência em termos computacional e de precisão de registro. O algoritmo de registro é muito simples, fácil de aplicar e razoavelmente eficiente em termos de complexidade computacional. Vários experimentos e medidas de desempenho são apresentados para ilustrar o desempenho do algoritmo.

## AUTOMATIC REGISTRATION BASED ON MULTIRESOLUTION DECOMPOSITION OF REMOTE SENSING IMAGES

#### **ABSTRACT**

Image registration is one of the basic image processing operations in remote sensing. With the increase in the number of images collected every day from different sensors, automated registration of multi-sensor, multi-spectral and multi-temporal images has become an important issue. In this thesis an efficient algorithm for the registration of multi-temporal images with similar spectral responses is proposed. The procedure is completely automatic and relies on the grey level information content of the images and their local wavelet transform modulus maxima. The algorithm is performed at progressively higher resolution, which allows for faster implementation and higher registration precision. The registration algorithm is very simple, easy to apply and reasonably efficient in terms of computational complexity. Extensive experimental results are provided to illustrate the performance of the algorithm as well as measurements of its performance.

## SUMÁRIO

	Pag.
LISTA DE FIGURAS LISTA DE TABELAS	
CAPÍTULO 1 INTRODUÇÃO	21
1.1 Proposta do Trabalho	23
1.2 Organização da Tese	26
CAPÍTULO 2 REGISTRO DE IMAGENS	29
2.1 O problema de Registro de Imagens	29
2.2 Técnicas de Registro	32
2.3 Transformações Espaciais	40
2.4 Algoritmos de Registro: estudo comparativo	44
CAPÍTULO 3 FUNDAMENTOS TEÓRICOS	49
3.1 Análise em Multi-resolução	49
3.2 Wavelets	<i></i> . 54
3.2.1 Transformada Wavelet	56
3.2.2 Transformada Wavelet Diádica Discreta	60
3.2.3 Representação em Wavelets para Sinais 1D	62
3.2.4 Representação em <i>Wavelets</i> para Imagens	64
CAPÍTULO 4 DETECÇÃO DE FEIÇÕES PONTUAIS	71
4.1 Detecção de Bordas através do Módulo dos Máximos da Tran	nsformada
Wavelet	73
4.2 Detecção de Feições Pontuais	75
4.3 Detalhes de Implementação	77
4.4 Resultados Experimentais	81

CAPÍTULO 5 MEDIDAS DE SIMILARIDADE: ANÁLISE DE	<u>}</u>
DESEMPENHO	91
5.1 Medidas de Similaridade	92
5.1.1 Coeficiente de Correlação	92
5.1.2 Medida de Similaridade baseada na Norma $L_1$	93
5.1.3 Detecção de Similaridade Sequencial	93
5.1.4 Operador de Casamento Morfológico: GHMT	94
5.2 Medidas de Confiança	97
5.3 Avaliação dos Critérios de Similaridade	99
5.3.1 Experimento 1: desempenho em função do tamanho da janela	100
5.3.2 Experimento 2: desempenho com medida de confiança	102
5.3.3 Experimento 3: desempenho dos métodos CORREL e L1G com medida	
de confiança	105
5.3.4 Conclusão	108
CAPÍTULO 6 ALGORITMO DE REGISTRO EM MÚLTIPLAS	}
RESOLUÇÕES	111
6.1 Pré-processamento	114
6.2 Detecção de Feições Pontuais	115
6.3 Transformação Espacial	116
6.4 Processo de Casamento	117
6.4.1 Casamento Inicial	117
6.4.2 Testes de Consistência	119
6.4.3 Refinamento dos pontos de controle iniciais	125
6.4.4 Algoritmo de Casamento Inicial	126
6.5 Processo de Refinamento do registro	126
6.5.1 Algoritmo de Refinamento	130

CAF	PÍTULO 7	RESULTADOS EXPERIMENTAIS	133
7.1	Experimento	A	133
7.2	Experimento	B	137
7.3	Experimento	C	142
7.4	Experimento	D	163
CAF	PÍTULO 8	AVALIAÇÃO DOS RESULTADOS	171
CAF	PÍTULO 9	CONCLUSÕES E DIREÇÕES PARA PESQUISAS	
		FUTURAS	189
9.1	Resumo e Co	nclusões	189
9.2	Futuras Pesq	uisas	191
REF	ERÊNCIAS	BIBLIOGRÁFICAS	193
APÊ	NDICE A		211
APÊ	NDICE B		213
APÊ	ENDICE C		217

### LISTA DE FIGURAS

Pág.

2.1	Registro de imagens	30
2.2	Registro baseado em área	33
3.1	Transformada de Fourier	50
3.2	Decomposição de um sinal $f$ em dois sinais aproximado $a$ e detalhe $d$	52
3.3	Pirâmide de imagens	53
3.4	Esquema de Codificação Sub-banda (SBC)	54
3.5	Funções seno e wavelet de Daubechies	55
3.6	Ilustração do cálculo da transformada wavelet	58
3.7	Decomposição em wavelets	64
3.8	Reconstrução de um sinal a partir de sua representação por wavelets	65
3.9	Disposição das imagens na representação wavelet	66
3.10	Decomposição de uma imagem $A^d_{2^j}f$	67
3.11	Reconstrução de uma imagem a partir de sua representação wavelet .	69
4.1	Esquema dos processos envolvidos na área de reconhecimento de padrões	71
4.2	Convenção adotada para o sistema de coordenadas $x$ e $y$ na represen-	
	tação da imagem digital $f$	78
4.3	Funções wavelet spline cúbica e a correspondente função de suavização	79
4.4	Vizinhos mais próximos do Ponto ${\cal P}$ na direção do vetor gradiente	82
4.5	Direções aproximadas indicadas pelo ângulo $lpha$	82
4.6	Imagem TM da cidade de Brasília	84
4.7	Imagem TM de uma região de agricultura	85
4.8	Imagem TM da região Amazônica	86
4.9	Detecção de feições pontuais para imagem da cidade de Brasília	87
4.10	Detecção de feições pontuais para a imagem da região de agricultura	88
4.11	Detecção de feições pontuais da imagem da região Amazônica	89
5.1	Aplicação do operador GHMT sobre um sinal unidimensional $f$	96

5.2	Desempenho dos métodos de correlação, L1G, DSS e GHMT em função	
	do tamanho da janela de casamento	101
5.3	Desempenho médio dos métodos CORREL, L1G, DSS e GHMT com	
	medidas de confiança	104
5.4	Taxa de alarmes falsos em função do número de casamentos corretos	
	para os métodos CORREL, L1G, DSS e GHMT	106
5.5	Desempenho dos métodos de correlação e L1G com medidas de con-	
	fiança	107
5.6	Taxa de alarmes falsos em função do número de casamentos corretos	
	para os métodos CORREL e L1G	109
6.1	Estágios de processamento envolvidos no algoritmo de registro	113
6.2	Esquema do processo de casamento inicial	118
6.3	Janelas de processamento correspondentes em diferentes níveis de re-	
	soluções	119
6.4	Casamento bidirecional	120
6.5	Teste de consistência	122
6.6	Processo de refinamento dos casamentos	127
7.1	Imagens usadas para simular as imagens distorcidas	136
7.1	Conclusão	137
7.2	Pontos de controle obtidos para corrigir a imagem EAS-1 distorcida .	138
7.3	Pontos de controle obtidos para corrigir a imagem SIR-C/X distorcida	138
7.4	Pontos de controle obtidos para corrigir a imagem distorcida do sensor	
	ótico acoplado ao balão	139
7.5	Pontos de controle obtidos para corrigir a imagem Lansdat-TM distor-	
	cida	139
7.6	Mosaico entre as imagens AES-1 original e simulada corrigida	140
7.7	Mosaico entre as imagens Landsat-TM original e a simulada, após a	
	correção	141
7.8	Pontos de controle obtidos para corrigir a imagem TM97 distorcida .	142
7.9	Composição colorida das imagens TM97 corrigida e TM95	143

7.10	Registro de imagens da Amazônia	148
7.10	Conclusão	149
7.11	Registro de imagens da Amazônia	150
7.11	Conclusão	151
7.12	Registro de imagens da Amazônia	152
7.12	Conclusão	153
7.13	Registro de imagens de área de agricultura	154
7.13	Conclusão	155
7.14	Registro de imagens (georeferenciada) de área de agricultura	156
7.15	Registro de imagens de área de agricultura	157
7.15	Conclusão	158
7.16	Registro de imagens de área urbana	159
7.16	Conclusão	160
7.17	Registro de imagens de área urbana	161
7.17	Conclusão	162
7.18	Registro das imagens JERS93 e JERS96	165
7.18	Conclusão	166
7.19	Registro das imagens JERS93 e JERS95	167
7.19	Conclusão	168
7.20	Registro das imagens JERS95B e JERS96B	169
7.20	Conclusão	170
8.1	Composição colorida das imagens TM945AM e TM925AM registrada	172
8.2	Composição colorida das imagens TM945AM e TM965AM registrada	173
8.3	Composição colorida das imagens TM945I e TM905I registrada	174
8.4	Composição colorida das imagens TM945RI e TM905I registrada $$	175
8.5	Composição colorida das imagens TM925AG e TM945AG registrada	176
8.6	Composição colorida das imagens TM944SP e SP953SP registrada .	177
8.7	Erros de registro (RMSE) para diferentes níveis de processamento	183
8.8	Erros de registro (RMSE) nos diferentes níveis de resolução, para as	
	imagens Landsat-TM5	184

8.9	Erros de registro (RMSE) nos diferentes níveis de resolução (L=6) para	
	as imagens JERS	

## LISTA DE TABELAS

		Pág.
2.1	Comparação dos métodos de Registro— Parte I	47
2.2	Comparação dos métodos de Registro— Parte II	48
4.1	Coeficientes dos filtros correspondentes a wavelet spline cúbica	80
7.1	Parâmetros do modelo de distorção (real) e estimados pelo método de	
	registro proposto	134
7.2	Imagens testes (óticas)	145
7.3	lmagens testes (radar)	163
8.1	Erros de registro (RMSE) e tempo de processamento	179
8.2	Erros de registro (RMSE) e tempo de processamento (L=0)	181
8.3	Parâmetros da função de transformação e erros (RMSE)	183
8.4	Erros (RMSE) para os diferentes níveis de resolução (L=6)	185

## CAPÍTULO 1

## INTRODUÇÃO

O registro de imagens é o processo de sobreposição de duas ou mais imagens de uma mesma área geográfica, de modo que pontos das imagens que correspondem aos mesmos objetos de uma cena coincidam espacialmente. Este processo de casamento ou sobreposição é realizado através de uma transformação espacial adequada, que relaciona as coordenadas espaciais das imagens e modela a distorção geométrica entre elas.

O processo de registro de imagens é fundamental em vários problemas práticos nas áreas de sensoriamento remoto, visão por computador, reconhecimento de padrões e análise de imagens médicas, tais como:

- integração de dados adquiridos a partir de sensores diferentes (fusão de imagens/sensores);
- análise de mudanças nas imagens adquiridas em diferentes épocas (registro temporal e detecção de mudanças);
- mosaico de imagens;
- identificação de um padrão de referência em uma imagem;
- determinação do movimento de um objeto.

Em sensoriamento remoto, há um grande número de sensores de recursos naturais com diferentes características geométricas e radiométricas. A combinação ou fusão de imagens adquiridas por estes sensores pode efetivamente melhorar o processo de extração de informação, sendo por isto uma técnica bastante utilizada. Por exemplo, a combinação de imagens dos satélites SPOT-HRV (Haute Resolution Visible) e Landsat-TM (Landsat-Thematic Mapper) tem sido muito usada no monitoramento do crescimento urbano. Imagens SPOT apresentam melhor resolução espacial que as imagens TM, enquanto estas apresentam melhor resolução espectral que as imagens

SPOT. A transformação IHS (Intensity-Hue-Saturation) pode ser usada para combinar a banda pancromática do SPOT com as bandas espectrais do TM gerando assim uma outra imagem colorida com alta resolução espacial (Carper et al., 1990; Vergara et al., 1996). Novas técnicas de fusão de imagens são apresenatadas em Brum (1989), Candeias (1992) e Mascarenhas et al. (1996). Nesta aplicação, é necessário que as imagens estejam devidamente registradas antes da fusão dos dados.

Um outro exemplo é a combinação de imagens óticas com imagens de radar. As imagens de radar são pouco afetadas por coberturas de nuvens, e além disso contêm informações que podem ser complementares na análise da região em estudo. Por exemplo, dados de radar de abertura sintética (SAR) obtidos a partir do sistema SIR-C (Shuttle Imaging Radar-C) e do satélite japonês JERS-1 (Japanese Earth Resources Satellite-1) combinados com imagens TM têm sido usados no mapeamento de inundações e vegetação na área de Manaus (Melack et al., 1994). Sensores SAR operam independentemente da iluminação solar e cobertura de nuvens, enquanto imagens TM fornecem informações adicionais na porção do espectro óptico. Como estes sensores estão em diferentes plataformas e em diferentes órbitas, o registro das imagens é um processo necessário antes da combinação e análise dos dados.

No campo da análise de imagens médicas, é procedimento comum a integração de informações obtidas por diferentes equipamentos tais como a tomografia computadorizada (CT), tomografia por emissão de pósitrons (PET) e imageamento por ressonância magnética (MRI) para detecção e localização de doenças (Maguire et al., 1985, 1991; Chen et al., 1997; Maes et al., 1997). Cada um destes equipamentos gera imagens com diferentes características geométricas, as quais também devem ser registradas antes de serem analisadas.

Na análise multi-temporal de imagens, o objetivo é detectar mudanças que tenham ocorrido em um determinado intervalo de tempo. Aplicações típicas da análise temporal de imagens são a detecção e monitoramento de mudanças e crescimento urbano (Foresti et al., 1980), monitoramento de desmatamento e diagnóstico médico no caso de imagens médicas. Um método simples de detectar estas modificações em

um par de imagens é sobrepô-las e identificar as diferenças entre elas. Como estas imagens são tomadas sob condições e em épocas diferentes, elas devem ser registradas antes do processo comparativo.

Em visão por computador, o registro é necessário, por exemplo, quando se quer determinar o movimento de um objeto, ou para se obter informação tridimensional de uma cena a partir de um par estereoscópico de imagens (Chapman e Mahani, 1989). Nesta categoria incluem-se, também, o problema de mosaico de imagens (Jaillon e Montanvert, 1994) e a medição da direção do vento em 3-D (Zheng e Chellapa, 1993). Um mosaico de imagens consiste em uma composição de imagens adquiridas de diferentes pontos de vista para construir uma imagem maior, dando uma visão global da cena.

O reconhecimento de um padrão de referência em uma imagem, como o reconhecimento de caracteres, verificação de assinatura e impressões digitais (Sasakawa et al., 1990), localização da posição e orientação de feições conhecidas tais como pistas de aeroporto, lotes de estacionamento, tanques de guerra, etc., também está relacionado com o problema de registro.

A diversidade de aplicações têm sido a causa para o desenvolvimento de inúmeros métodos de registro de imagens com vários graus de complexidade. O fato de muita pesquisa sobre este assunto ainda estar em desenvolvimento é indicativo da complexidade do problema e da necessidade dos usuários de imagens de sensores remotos de terem métodos eficientes e adequados aos seus dados e aplicações.

#### 1.1 Proposta do Trabalho

Existe um grande número de algoritmos de registro que podem ser aplicados em casos específicos e que produzem resultados úteis, dentro de suas limitações. As pesquisas estão caminhando para o desenvolvimento de metodologias que melhorem métodos já existentes, que combinem diferentes algoritmos dentro de um mesmo sistema (Rignot et al., 1991; Moigne et al., 1997; Dare et al, 1997) ou que apresentem soluções para os problemas que dificultam o desenvolvimento de um sistema de registro automático

ou semi-automático. O tipo de dados e o de aplicações continuam a ser pontos importantes a serem levados em conta no desenvolvimento de tais sistemas.

Agências do mundo todo tem lançado ou estão preparando para lançar vários sistemas de obervação da terra, de alta resolução, muti-bandas e multi-sensores incluindo SPOT-3, Landsat-7, ENVISAT 1, TRMM (Tropical Rainfall Measuring Mission), EOS-AM (Earth Observing System) e o satélite brasileiro SSR (Satélite de Sensoriamento Remoto). A alta resolução significa que as correções geométricas e o registro destas imagens devem ser cada vez mais precisos. O aumento de volume de dados implica na necessidade de desenvolvimento de métodos de registro de imagens automáticos ou semi-automáticos com boa eficiência computacional.

Métodos automáticos baseados em feições tais como o proposto por Li et al. (1995), embora adequados para imagens multi-sensores não são adequados para o registro de imagens multi-temporais. Nas imagens multi-temporais a serem registradas, os contornos correspondentes a um mesmo objeto pode apresentar diferenças devido às mudanças sazonais, o que inviabilizam o uso destes métodos. A obtenção dos contornos dos objetos torna-se ainda mais difícil quando as imagens possuem forte característica textural, exigindo assim técnicas de extração de feições mais sofisticadas.

O processo de registro de imagens pode ser demorado, difícil de ser efetuado e de baixa precisão se o procedimento não é bem entendido pelo usuário.

Levando em conta todos os aspectos mencionados acima, propomos neste trabalho o desenvolvimento de uma metodologia de registro que seja capaz de registrar imagens multi-temporais de mesmo sensor ou sensores diferentes com respostas espectrais similares, de fácil utilização pelo usuário e computacionalmente eficiente. Os principais problemas que consideramos são a identificação e casamento automáticos de feições nas imagens, que são as duas fases mais complexas e que demandam mais tempo em todo o processo de registro. A idéia é reduzir ao mínimo a interação do usuário nas tarefas de identificação e casamento de feições.

A característica principal do nosso algoritmo é a integração das fases de detecção de feições pontuais, casamento destas feições e verificação da consistência de casamentos dentro de um esquema de multi-resolução através da transformada wavelet, que permite as seguintes vantagens:

- efetua um registro confiável em um nível de resolução mais baixo, e então refina o resultado do registro nos níveis de resolução mais altos;
- apresenta boa eficiência computacional, obtida pela redução do tamanho das imagens de baixa resolução;
- é de implementação relativamente simples;
- decompõe a imagem em sub-imagens que possuem características distintas,
   e que são úteis na detecção de feições, redução de ruídos, etapas importantes
   no processo de registro de imagens.

Para se obter as feições usadas no processo de casamento, propomos um método de extração de feições baseado nos máximos locais do módulo da transformada wavelet. A nossa contribuição, é a automatização do procedimento que, basicamente com dois parâmetros, gera um conjunto de feições significativas e consistentes em todos os níveis de resolução da decomposição wavelet e no nível 0 (imagem original).

Realiza-se também uma avaliação de desempenho de quatro medidas de similaridade, a saber, o coeficiente de correlação, a detecção de similaridade seqüencial, uma medida baseada na norma  $L_1$  e um operador de casamento morfológico. O resultado desta avaliação aponta o coeficiente de correlação como a medida de melhor desempenho e, neste trabalho, ela é usada para o casamento automático das feições extraídas nas imagens.

O processo de verificação de consistência dos casamentos, para eliminar os casamentos falsos, é uma fase muito importante no processo de registro. A escolha errada dos pontos de controle inviabiliza o registro correto. No caso de uma transformação rígida pode-se usar a hipótese de que as distâncias euclidianas entre dois pontos de

controle nas imagens são preservadas e assim eliminar os pontos casados que não atendem esta condição. Para relaxar esta restrição, propomos um novo algoritmo que usa uma busca combinatória entre os N casamentos, obtidos na fase inicial, para selecionar aqueles corretos. Embora este método exija um maior esforço computacional, ele é independente do modelo de distorção das imagens.

A principal restrição do algoritmo de registro proposto é que ele é mais apropriado para registrar imagens com respostas espectrais similares, pois ele utiliza a informação espectral no processo de casamento de feições. Além disso, o ângulo de rotação e a escala entre as imagens devem ser pequenos. O problema de escala é resolvido através da reamostragem das imagens para um mesmo tamanho de pixel. O problema da rotação, para as imagens de satélites, é irrelevante posto que a diferença de rotação entre as imagens, normalmente, é pequena ou é previamente conhecida.

O algoritmo proposto é fácil de se aplicar porque ele precisa basicamente de dois parâmetros: um que define o número de níveis da decomposição wavelet e outro que define o número de feições extraídas usadas para gerar os pontos de controle. Em termos de complexidade computacional o algoritmo é eficiente. Comparado ao processamento realizado no nível de resolução 0, o processamento em múltiplas resoluções chega a ser dezenas de vezes mais rápido, com a mesma ou melhor precisão de registro.

A sua robustez e bom desempenho é comprovado através de vários testes com imagens óticas e de radares.

#### 1.2 Organização da Tese

Este trabalho é organizado da seguinte forma. No Capítulo 2 é apresentado, de modo resumido, o problema de registro de imagens e são discutidos alguns trabalhos sobre registro automático de imagens recentemente propostos na literatura. Neste Capítulo, é ainda apresentado um estudo comparativo entre estes métodos, que é uma boa referência para os leitores interessados no problema de registro de imagens.

Como a transformada wavelet é utilizada como uma importante ferramenta neste trabalho, e por ser uma técnica relativamente recente na área de processamento de imagens, no Capítulo 3 é apresentado um resumo da teoria correspondente, fazendo um paralelo entre a formulação matemática e a implementação prática.

No Capítulo 4 a relação entre a transformada wavelet e o problema de detecção de bordas, formalizada por Mallat e Zhong (1992b), é apresentada. Um algoritmo automático de detecção de feições pontuais basedo na transformada wavelet, é também apresentado neste Capítulo.

Um estudo do desempenho de quatro medidas de similaridade é apresentado no Capítulo 5.

No Capítulo 6 é apresentado o método de registro proposto, onde cada fase do algoritmo de registro é descrito em detalhes.

Para ilustrar o desempenho do algoritmo de registro proposto, no Capítulo 7 apresentados vários resultados experimentais. O método proposto são aplicado para registrar alguns pares de imagens multi-temporais dos satélites Landsat-TM, SPOT, SIR-C e JERS. A avaliação dos resultados é realizada através do erro de registro, medido pela raiz do erro médio quadrático (RMSE) e também através da análise visual. As imagens registradas são sobrepostas e a precisão do casamento entre elas é verificada visualmente.

Finalmente, no Capítulo 8 apresentam-se um resumo do trabalho, conclusões e propostas para pesquisas futuras.

### CAPÍTULO 2

#### REGISTRO DE IMAGENS

### 2.1 O problema de Registro de Imagens

Imagens adquiridas por sensores a bordo de satélites ou aeronaves contém distorções radiométricas e geométricas. As distorções radiométricas afetam os valores dos níveis de cinza da imagem, e são provocadas por fatores atmosféricos e erros de medição dos instrumentos. As distorções geométricas mudam a geometria da imagem de modo que a posição, o tamanho e a forma do pixel são alterados durante o processo de aquisição de imagens. São várias as fontes de erros geométricos tais como a rotação da Terra durante o processo de aquisição, a curvatura da Terra, variações da velocidade e altitude da plataforma, mudanças na elevação topográfica, etc. (Richards, 1986). Distorções sistemáticas, como aquelas provocadas pela rotação da Terra e órbita não-polar do satélite e variações de atitude e altitude, podem ser corrigidas usando modelos orbitais e dados de calibração (registro paramétrico). Outras distorções imprevisíveis devem ser removidas com a ajuda de pontos de controle e modelos matemáticos de distorção, que é o procedimento usado no processo de registro de imagens.

O problema de registro de duas imagens obtidas de uma mesma cena pode ser entendido de uma maneira simples, como a operação que faz com que uma imagem se sobreponha perfeitamente a outra, de modo que cada par de pontos sobrepostos corresponda ao mesmo ponto da cena real (Erthal, 1985).

Dois tipos de registro podem ser realizados: 1) mapa-imagem e 2) imagem-imagem. No registro mapa-imagem, a imagem é colocada dentro de um sistema de coordenadas planas de um certo sistema de projeção cartográfica, tal como o *Universal Transverse Mercator* (UTM) ou o *Polar Stereographic* (PS). Neste caso dizemos que a imagem está geo-referenciada ou geo-codificada (Niblack, 1986).

No registro imagem-imagem, uma imagem é registrada com outra imagem da mesma

área geográfica, de diferente ou mesmo sensor, ou de diferente data. A imagem considerada como o padrão de referência, sobre a qual a segunda é sobreposta, é chamada de imagem de referência. A imagem a ser registrada com a imagem de referência é denominada imagem de ajuste ou imagem de entrada. A Figura 2.1 ilustra o processo de registro imagem-imagem.

Na prática, a imagem é registrada com um mapa dentro de um sistema de projeção cartográfica, e esta imagem registrada é tomada como a imagem de referência sobre a qual outras imagens são registradas. Com este procedimento, maior precisão de registro pode ser obtido em relação aquele que registra cada imagem com o mapa independentemente (Niblack, 1986).

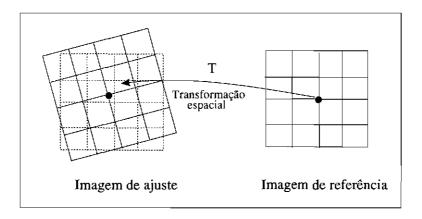


Fig. 2.1 - Registro de imagens.

O problema central do processo de registro é determinar a função de mapeamento, que modela a distorção entre os dados a serem registrados, o que pode ser realizado de três formas:

- usando o modelo de distorção geométrica produzido pelo sistema sensor (modelo da cena/sensor);
- 2) usando pontos de controle para estimar o modelo ou,
- 3) combinando os pontos de controle e o modelo da cena/sensor.

O primeiro método envolve um modelo detalhado da deformação, que muitas vezes não é fácil de se conseguir. Conhecido este modelo, o terceiro método é o mais preciso. O modelo cena/sensor é usado para corrigir as distorções da imagem e os pontos de controle são usados para refinar o registro. Desta forma, pode-se obter uma precisão de registro com erros muitos pequenos, podendo chegar a 0,15 pixel (Niblack, 1986).

O segundo método é o mais simples e por isso o mais usado. Ele pode ser realizado em quatro fases:

- 1) identificação de feições nas imagens a serem registradas;
- 2) casamento das feições;
- 3) determinação da transformação espacial;
- 4) interpolação.

O processo de identificação de feições identifica um conjunto de feições (objetos) relevantes nas duas imagens, tais como intersecções de linhas, contornos de regiões, regiões, pontos de borda, etc. Estas feições devem estar presentes e, de preferência, bem delineadas em ambas as imagens.

O processo de casamento estabelece a correspondência entre as feições. Cada feição na imagem de ajuste é casada com a correspondente feição na imagem de referência. As feições, assim casadas, são identificadas por coordenadas espaciais que identificam as suas posições nas imagens. Estas coordenadas espaciais são usualmente chamadas de pontos de controle. O processo de casamento envolve a busca de correspondências entre um grande número de feições, e portanto, é uma tarefa computacionalmente cara.

Após o processo de casamento temos um conjunto de pares de pontos de controle. Este conjunto é usado para determinar a função de transformação, que modela matematicamente a distorção entre as imagens de referência e de ajuste, a qual é aplicada à grade da imagem de ajuste.

O processo de interpolação é bastante simples, e as técnicas do Vizinho Mais Próximo (VMP), Bilinear e Convolução Cúbica são as mais utilizadas (Niblack, 1986, p.142).

Uma técnica de interpolação híbrida (bilinear e VMP) é proposta por Câmara (1983).

#### 2.2 Técnicas de Registro

A tarefa mais difícil, no processo de registro de imagens, é a identificação e o casamento de feições para obter os pontos de controle. Esta tarefa é crucial para a precisão do registro das imagens.

O método de registro manual é o mais simples e o mais utilizado nas diversas aplicações da área de sensoriamento remoto. Neste método, as fases de identificação e casamento de feições nas imagens são realizadas simultaneamente, e com a assistência de um intérprete. As imagens são visualizadas no monitor do computador e o usuário escolhe as feições correspondentes que aparecem nas duas imagens. Possíveis candidatos incluem lagos, rios, cruzamento de ruas, pontes, ou qualquer outro objeto que seja claramente visível nas duas imagens. A cada objeto identificado é designada uma coordenada na imagem, e os pares de coordenadas correspondentes são tomados como pontos de controle.

Para se obter um registro de boa precisão, um grande número de pontos de controle, bem distribuídos em toda a extensão da imagem, deve ser selecionado. Isto é uma tarefa muito tediosa e repetitiva, e muitas vezes o processo não é eficiente quando o volume de dados é grande. Então, há a necessidade de técnicas de registro semi-automático ou automático que requeiram pouca ou nenhuma supervisão humana. Segundo a natureza das feições utilizadas, estas técnicas podem ser agrupadas em duas categorias: técnicas baseadas em área e técnicas baseadas em feições.

Nos métodos baseados em área, uma pequena janela de pontos na imagem de referência é estatisticamente comparada com janelas de mesmo tamanho na imagem de ajuste. Este processo é ilustrado na Figura 2.2. Considere a imagem de ajuste S com M linhas e N colunas, e n janelas  $W_z$ , z=1,...,n, com K (inteiro e ímpar) linhas e L (inteiro e ímpar) colunas extraídas da imagem de referência R e centradas nos pontos  $(a_z, b_z)$ . Seja  $S_{ij}$  uma subimagem de S com sua coordenada superior e

esquerda representada por (i, j), onde

$$S_{ij}(l,m) = S(i+l,j+m)$$
, (2.1)

para  $0 \le l \le K-1$ ,  $0 \le m \le L-1$  e,  $0 \le i \le M-K$ ,  $0 \le j \le N-L$ . Cada janela  $W_z$  é comparada com todas as subimagens  $S_{ij}$  na imagem S. Depois de encontrar a subimagem  $S_{ij}$  que melhor se casa com  $W_z$ , seus centros  $(a_z,b_z)$  e (i+(K-1)/2,j+(L-1)/2) são tomados como pontos de controle. Então, esses pontos de controle podem ser usados para calcular os parâmetros da função de mapeamento.

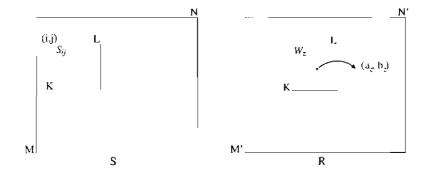


Fig. 2.2 - Registro bascado em área.

O processo de comparação usa medidas de similaridade entre as duas janelas. As técnicas mais comuns utilizam a correlação normalizada, o coeficiente de correlação, e a detecção de similaridade sequencial (Barnea e Silverman, 1972; Pratt, 1974).

Vários trabalhos têm sido publicados (Mascarenhas e Pereira, 1982, 1983a,b; Erthal et al., 1983; Mascarenhas e Erthal, 1984, 1992; Erthal, 1984; Câmara e Ii, 1985; Cideciyan et al., 1992; Zheng and Chellapa, 1993; Kher e Mitra, 1993; Banon e Faria, 1997). Por muito tempo, os métodos de casamento baseado em área foram limitados ao problema de registro onde as imagens eram desalinhadas por uma pequena diferença de rotação. Zheng and Chellapa (1993) propuseram uma nova solução em que a rotação é estimada e corrigida primeiramente. Eles usam a técnica shape-from-shading para estimar as direções da iluminação na imagem. Tomandose as diferenças entre estas iluminações o ângulo de rotação entre as imagens é

calculado. As feições nas imagens são extraídas através da transformação de Gabor (Manjunath et al., 1992) e casadas umas às outras usando a correlação no domínio do espaço. No método de Zheng and Chellapa (1993), o método de estimação do ângulo de rotação funciona bem na maioria dos casos. Entretanto, se uma cena inclui muitos objetos e construções o método tem grandes chances de falhar se as condições de iluminação nas imagens não são equivalentes.

Os métodos de correlação podem ser eficientemente implementados no domínio da transformada de Fourier usando a transformada de Fourier rápida (FFT). Além disso, algumas propriedades da transformada de Fourier podem ser usadas para adquirir invariância a translação, rotação e escala (Anuta, 1970; Granlud, 1972; Kuglin e Hines, 1975; Persoon e Fu, 1977; Jensen at al, 1987; Decastro e Morandi, 1987). Por exemplo, se duas imagens se diferem apenas por um deslocamento, a diferença de fase da transformada de Fourier das duas imagens pode ser usada para estimar este deslocamento. Esta técnica é mais adequada para a correção de imagens que tenham sofrido apenas distorção rígida e que tenham grande sobreposição.

Trabalhos recentes, usando técnicas de Fourier, têm sido desenvolvidos para correção tanto do efeito de translação, como de rotação e escala (Cideciyan et al., 1992; Kher e Mitra, 1993). Cideciyan et al. (1992) usam algumas propriedades da transformada de Fourier para separar o efeito de translação dos efeitos de rotação e escala. Esses métodos são mais adequados na correção de imagens que tenham sofrido apenas distorção rígida. Outras restrições deste método são que a diferença de rotação e escala entre as imagens seja pequena, e que haja uma grande sobreposição entre as imagens de tal forma que os objetos de interesse estejam completamente contidos em ambas as imagens.

As principais dificuldades encontradas no uso dos métodos baseados em área ocorrem quando:

- as imagens são tomadas de pontos de vista extremamente diferentes;
- as imagens são tomadas de uma cena que não contem textura visual ade-

quada, e

• as imagens possuem grandes diferenças espectrais.

No método de registro baseado em feições duas tarefas são envolvidas no processo de identificação dos pontos de controle: extração de feições e casamento de feições. Na fase de extração de feições a imagem é representada, em uma forma compacta, por um conjunto de feições e seus atributos. O processo de casamento é realizado neste mesmo espaço de representação. A imagem, neste caso, pode ser representada no domínio do espaço ou no domínio da transformada.

No domínio do espaço, as feições mais comumente usadas são bordas, regiões, linhas, intersecções de linhas, centróides de regiões, pontos de descontinuidades na curvatura, etc. Os contornos das regiões e bordas têm sido frequentemente usados como feições. Estas são extraídas por técnicas de segmentação e detecção de bordas tais como o operador de Canny (Canny, 1986), operador Laplaciano da Gaussiana (Pratt, 1991) e crescimento de região (Ballard e Brown, 1982; Bins et al., 1996).

Geralmente os algoritmos de casamento de feições usam atributos de forma (perímetro, momentos invariantes à translação, alongamento), cor, textura, relações definidas pelo arranjo espacial. Normalmente, procura-se usar propriedades invariantes à rotação, escala, e variações nos níveis de cinza da imagem. Cada objeto na primeira imagem é comparado com os objetos candidatos em potencial ao casamento na segunda imagem, usando os seus atributos. Um par de objetos que possui atributos similares são aceitos como casados.

Vários trabalhos têm sido propostos nesta categoria (Goshtasby e Stockman, 1985, Goshtasby et al. 1986; Câmara, 1986; Bins, 1988; Ton e Jain, 1989; Shenk et al., 1991; Toth e Schenk, 1992; Dana e Anandan, 1993; Flusser e Suk, (1994) e Abbasi-Dezfouli e Freeman, 1994; Li et al., 1995). Flusser e Suk (1994) representam cada região por atributos de forma, que são usados no casamento destas regiões. Goshtasby et al. (1986) e Ton e Jain (1989) usam os centróides das regiões como pontos de controle, e técnicas de agregamento e relaxação para o casamento dos

objetos. Goshtasby et al. (1986) observam, ainda, que para imagens com distorções geométricas significativas, feições tais como intersecções de linhas e finais de linha deveriam ser usadas ao invés de centróides. Li et al. (1995) usam, além dos centróides de regiões fechadas, pontos salientes dos contornos abertos como pontos de controle. Os contornos abertos e fechados são representados por códigos de cadeia e por um conjunto de atributos de forma. Estes atributos e a medida de correlação do código de cadeia são usados para casar os contornos. Schenk et al. (1991) e Toth e Schenk (1992) propõem um método similar ao de Li et al. (1993) usando outra forma de representação dos contornos. Abbasi-Dezfouli e Freeman (1994) extraem regiões homogêneas e usam seus atributos tais como área, perímetro, concavidade, etc, para casá-las.

Dana e Anandan (1993) usam os contornos dos objetos, obtidos por um detector de bordas em diferentes escalas, para determinar a transformação afim entre as imagens. Os parâmetros da transformação são refinados, de forma iterativa, nos diversos níveis de resolução usando um método baseado no gradiente. Ventura et al. (1990) usam regras lógicas no processo de casamento dos objetos. Para cada par de objetos nas imagens, as diferenças dos valores dos atributos são calculadas e classificadas dentro de um conjunto de rótulos (pequena, muito pequena, etc.). Então uma Árvore Lógica Multi-Valor (MVLT) decide se um par de objetos é "similar", "diferente", ou "quase-similar" baseado nos valores dos rótulos.

No domínio da transformada, a imagem pode ser decomposta e representada por um conjunto de coeficientes. Por exemplo, quando a transformada de Fourier é usada para representar a imagem no domínio da freqüência, a informação de borda na imagem pode ser adquirida considerando o conteúdo das altas freqüências no domínio de Fourier. Entretanto, o problema é que a transformada de Fourier dá uma informação global da distribuição das freqüências, mas não tem informação de localização espacial necessária para realizar o casamento no domínio espacial. Transformações espaciais tais como a transformada de wavelet (WT) são mais adequadas a esta aplicação (Mallat, 1989a,b; Chui, 1992).

Nesta categoria vários métodos têm sido desenvolvidos. Alguns mais recentes são os trabalhos de Chandra (1992), Djamdji et al. (1993, 1995), Moigne (1994), Corvi e Nicchiotti (1995), Li e at. (1995) e Deubler e Olivo (1997). Chandra (1992) usa a transformada de Haar para extrair um conjunto de feições (pontos). Janelas nas imagens são representadas por um número pré-determinado dos coeficientes da transformada de Haar de maiores valores, os quais são correlacionados para determinar a correspondência entre os pontos. Moigne (1994), Corvi e Nicchiotti (1995), e Djamdji et al. (1993) selecionam os coeficientes da transformada wavelet mais significativos para representar as feições nas imagens. Moigne (1994) calcula o ângulo de rotação entre as imagens através da correlação no domínio da transformada. Corvi e e Nicchiotti (1995) usam a técnica de agregamento para estabelecer a correspondência entre os pontos. Deubler e Olivo (1997) usam a sensibilidade à orientação das diferentes sub-bandas da transformada wavelet para determinar o deslocamento entre duas imagens que apresentam deformação global e local.

Embora os métodos baseados em feições sejam mais adequados para o registro de imagens adquiridas por diferentes sensores, algumas limitações são impostas. A maioria dos algoritmos de casamento de bordas se baseia no fato das bordas aparecerem bem definidas e espalhadas nas imagens, o que nem sempre acontece na prática. No caso de imagens com áreas densamente texturizadas ou moderado nível de ruído, estes algoritmos também podem se confundir e falhar, justamente onde os métodos baseados em área funcionam melhor. Além disso, o método exige técnicas de processamento sofisticadas no processo de extração e casamento de feições. Assim, o operador tem que decidir quais algoritmos devem ser usados em cada estágio, o que coloca mais empecilhos à automação.

De um modo geral, os algoritmos de registro são diferentes no sentido de que eles combinam técnicas diferentes para identificação e casamento de feições e determinação da função de mapeamento.

A despeito da variedade de técnicas que têm sido usadas para solucionar o problema, o registro de imagens de radar com imagens óticas ou de radar é um tema que vem recebendo muita atenção embora somente resultados parciais tenha sido obtidos. Este problema é relativamente complexo porque as imagens de radar apresentam (Dare e Dowman, 1996):

- diferentes características geométricas;
- ruído speckle;
- diferenças no brilho relativo entre os correspondentes objetos nas imagens;
- distorções geométricas relacionadas com o terreno. A imagem de radar de uma região montanhosa nem sempre reflete as variações radiométricas do terreno mas sim as variações geométricas.

Uma abordagem que tem sido proposta na literatura (Rignot et al., 1991; Takeuchi, 1993) é o uso de dados auxilares no processo de registro tais como modelo digital de elevação do terreno (DEM). Tais dados, quando disponíveis, podem fornecer informações complementares e auxiliar no processo de registro, gerando assim melhores resultados (Toutin, 1995).

Guindon (1985) usa imagens de radar simuladas, a partir de modelos digitais de terreno (DTM), para compensar os efeitos de relevo. A restrição é que nem sempre se tem o dado DTM e, quando o tem, nem sempre é preciso. Um trabalho mais recente é o de Dare e Dowman (1996), que desenvolveram um sistema automático para registrar imagens SAR e SPOT. Eles usam informações da órbita do sensor SAR e DEM para geocodificar (remover erros de perspectiva e erros devido ao terreno) automaticamente a imagem SAR com alto grau de precisão. Dare e Dowman (1996) comparam algumas técnicas de filtragem para redução do ruído speckle e propõem uma metodologia para extrair feições de imagens SAR, as quais são usadas no processo de casamento de feições.

Wu e Maitre (1990) propõem um método para registrar imagens SAR e óticas. Eles usam linhas da costa marítima, que possuem alto contraste em ambas as imagens, como feições a serem casadas. Pontos de inflexão ao longo dos contornos são detectados e o casamento é feito através de uma busca combinatória.

De forma diferente, Li et al. (1995) usam um modelo de contorno ativo (Kass et al., 1987; Williams e Shah, 1992), que consiste em usar os contornos da imagem ótica como condição inicial na detecção dos contornos na imagem de radar. O uso desta técnica é limitado ao registro de imagens com pequenas diferenças de deslocamentos (5 pixels). Kher e Mitra (1993) têm realizado o registro de imagens SAR no domínio da Transformada de Fourier. Antes de aplicar o algoritmo de registro, as imagens são processadas por um filtro morfológico para reduzir o efeito do ruído *speckle*.

Outro tópico que merece ser mencionado é o registro automático de mapas com imagens. Este é um problema complexo e pouca pesquisa se tem feito nesta direção. Três aspectos devem ser levados em conta na abordagem deste problema: 1) a escolha do formato dos dados: vetorial ou matricial; 2) escolha da escala do mapa; 3) remoção de símbolos e textos no mapa. Caves et al. (1991) apresentam uma metodologia para registrar uma imagem SAR, obtida de uma região sem muito relevo, com um mapa. Eles usam a técnica de correlação para casar as junções na imagem e no mapa, mas os resultados não são muito satisfatórios devido ao problema do ruído speckle. Ventura et al. (1990) têm também testado o seu método para registrar mapa com imagem TM, com o auxílio do modelo de elevação do terreno e outras informações relativas ao processo de aquisição dos dados e natureza da região em estudo. Morgado e Dowman (1997) registraram fotografias aéreas e mapas. As feições (bordas) são casadas usando como atributos o tamanho, forma e perímetro dos objetos. O registro grosseiro é refinado usando uma técnica de programação dinâmica. A imagem é registrada com o mapa usando uma transformação afim, cujos parâmetros são calculados pelo método dos mínimos quadrados. Bartl et al. (1996) identificam feições de bordas em imagens Landsat-TM, as quais são casadas com um mapa cadastral através de um algoritmo de casamento baseado na relaxação probablilística.

#### 2.3 Transformações Espaciais

Uma transformação espacial é uma função (mapeamento) que estabelece a correspondência espacial entre todos os pontos de uma imagem de entrada e a imagem de saída. No processo de registro de imagens, esta função modela matematicamente a distorção da imagem de entrada (imagem de ajuste) em relação à imagem de saída (imagem de referência). Nos métodos tradicionais de registro, a função de transformação é determinada usando os pontos de controle. A escolha adequada da função de transformação é de fundamental importância para a qualidade do registro.

Dados N pares de pontos de controle  $\{[(X_i, Y_i), (x_i, y_i)], i = 1, ..., N\}$  nas duas imagens de uma mesma área geográfica, a função de transformação pode ser formulada da seguinte maneira:

$$x_i = T_1(X_i, Y_i)$$

$$y_i = T_2(X_i, Y_i)$$
(2.2)

onde  $(X_i, Y_i)$  e  $(x_i, y_i)$  correspondem às coordenadas dos pontos na imagem de referência e na imagem de ajuste, respectivamente. As funções  $T_1$  e  $T_2$  especificam a transformação espacial e mapeiam as coordenadas dos pontos da imagem de referência para as coordenadas dos correspondentes pontos na imagem de ajuste. Este tipo de mapeamento é conhecido por mapeamento inverso, porque mapeia a imagem de saída sobre a imagem de entrada (Wolberg, 1990). Para cada posição (X,Y) do pixel de saída, calcula-se as coordenadas (x,y) do pixel da imagem de ajuste e atribui-se o seu correspondente valor de nível de cinza para a posição (X,Y) na imagem de saída. Geralmente os valores das coordenadas (x,y) não são números inteiros, e então deve-se usar um interpolador para estimar o valor do nível de cinza do pixel a partir dos níveis de cinza de seus pixels vizinhos na imagem de ajuste.

As funções de transformação usadas para alinhar duas imagens podem ser globais ou locais. Uma transformação global é dada por uma única equação que registra otimamente todos os pixels nas duas imagens. Na transformação local há várias

equações, e cada uma delas é usada para uma determinada região da imagem. As transformações locais, geralmente, são mais precisas mas exigem maior esforço computacional.

As transformações globais mais frequentemente usadas na área de sensoriamento remoto são as transformações rígidas, afins e polinomiais. Uma transformação rígida é aquela em que a forma e o tamanho relativo dos objetos nas imagens são preservados após a aplicação da transformação. Uma classe de transformação rígida é a transformação de similaridade, composta da combinação de uma rotação  $(\theta)$ , translação nas direções x e y  $(\Delta x, \Delta y)$  e uma mudança de escala (s). Esta transformação pode ser representada da seguinte forma:

$$x = s[\cos(\theta)X + \sin(\theta)Y] + \Delta X$$
  

$$y = s[-\sin(\theta)X + \cos(\theta)Y] + \Delta Y,$$
(2.3)

onde (x,y) e (X,Y) são pontos de controle correspondentes nas imagens de ajuste e de referência, respectivamente.

A transformação de similaridade é suficiente para registrar duas imagens adquiridas sob o mesmo ângulo de visada mas em posições diferentes. Ou seja, a câmera pode ser movida e rotacionada em torno do seu eixo óptico. Neste caso, um objeto quadrado na imagem é transformado em outro quadrado de tamanho diferente devido ao parâmetro escala.

Quando consideramos diferenças de escala nas direções x e y, temos uma função com cinco parâmetros  $(s_x, s_y, \theta, \Delta x, \Delta y)$ , à qual denominamos de função afim ortogonal:

$$x = s_X cos(\theta) X + s_Y sin(\theta) Y + \Delta X$$
  

$$y = -s_X sin(\theta) X + s_Y cos(\theta) Y + \Delta Y.$$
 (2.4)

Nesta transformação um quadrado se transforma em um retângulo, devido aos parâmetros de escalas em x e y.

A transformação afim, propriamente dita, ou função afim genérica leva em conta a distorção de cizalhamento e diferenças de escala nas duas direções x e y e pode ser representada por:

$$x = a_1X + a_2Y + a_3$$
  
 $y = b_1X + b_2Y + b_3,$  (2.5)

onde  $a_i$  e  $b_i$ , i=1,2,3 são os parâmetros da função a serem determinados. Ângulos e tamanhos dos objetos na imagem original não são mais preservados, mas linhas paralelas permanecem paralelas após o registro. Por exemplo, um quadrado na imagem original se transforma em um paralelograma na imagem registrada devido ao ângulo residual.

A transformação afim é muito usada em aplicações na área de sensoriamentro remoto e é uma boa aproximação para imagens tomadas sob condições de imageamento similares (Dana e Anandan, 1993), e que têm sido geometricamente corrigidas (por exemplo, em relação a rotação e curvatura da Terra) (Donze et al, 1989).

Outras transformações tais como a projetiva e perspectiva são usadas também nos casos de distorções espaciais mais genéricas (Brown, 1992). As transformações projetiva e perspectiva consideram distorções provocadas pela projeção dos objetos da cena no plano de imagem (visada lateral).

No caso de não se ter muita informação sobre a transformação necessária para alinhar as imagens, as transformações polinomiais são mais frequentemente usadas. Elas podem ser representadas por duas funções de transformação espacial da forma:

$$x = \sum_{i=0}^{m} \sum_{j=0}^{i} a_{ij} X^{i} Y^{i-j}$$

$$y = \sum_{i=0}^{m} \sum_{j=0}^{i} b_{ij} X^{i} Y^{i-j},$$
(2.6)

onde  $a_{ij}$  e  $b_{ij}$  são os coeficientes dos polinômios a serem determinados e m é o grau ou ordem do polinômio.

O modelo polinomial pode ser usado quando as imagens são bem construídas, como aquelas geradas por sensores a bordo de satélites, onde geralmente as plataformas têm condições de vôo e atitude estáveis. Neste caso, uma única função de transformação é suficiente para modelar a distorção (Buiten e Putten, 1997).

Este modelo é suficientemente geral para poder representar vários tipos de transformações geométricas. Assim o modelo polinomial de ordem 1 pode representar uma transformação consistindo de rotação, escala e translação dos eixos. O modelo polinomial de ordem 2 é apropriado para representar, por exemplo, as perspectivas. Polinômios de grau maior do que três devem ser usados com muita precaução devido ao comportamento imprevisível das funções polinomiais de alto grau (Richards, 1986). Estas geram resultados com boa precisão na vizinhança dos pontos de controle, mas podem gerar distorções em regiões da imagem fora desta vizinhança. Na prática é aconselhado usar o menor grau m, suficiente para modelar as distorções geométricas da imagem, e usar um número grande de pontos de controle com máxima precisão, descartando aqueles pares duvidosos.

Para contornar o problema de distorções geométricas locais, várias funções de transformação têm sido propostas (Goshtasby, 1986, 1987, 1988a, 1988b; Flusser, 1992). Estas funções são caracterizadas por usarem somente os pontos de controle vizinhos ao ponto que se deseja estimar, ao invés de usarem todos os pontos como no método global.

A função de mapeamento de superfície spline (Burrough, 1987) gera bons resultados no caso de registro de imagens com distorções locais. Entretanto, devido a sua complexidade computacional, ela não é muito adequada em aplicações práticas. Flusser (1992) propõe um método para determinação de uma função sensível a distorções geométricas não-lineares, com precisão comparável mas muito mais rápida que as funções spline.

#### 2.4 Algoritmos de Registro: estudo comparativo

A diversidade dos parâmetros do sistema de imageamento, das regiões imageadas, tamanho das áreas de estudo, estratégias de computação, e aplicações torna quase impossível a comparação quantitativa dos métodos de registro de imagens (Loew, 1997). Tendo isto em mente, procuramos realizar um estudo comparativo entre alguns algoritmos de registro de imagens recentes, segundo os seguintes tipos de informação:

- Sensores usados nos experimentos;
- Feições usadas no processo de casamento;
- Técnicas de casamento usadas;
- Número de experimentos usados na avaliação do método;
- Tipo de transformação espacial;
- Limitações e méritos do método.

O objetivo desta comparação é caracterizar os métodos de registro e assim, facilitar a escolha do método mais adequado a uma dada aplicação.

A escolha dos algoritmos mecionados nesta seção foi ditada, primariamente, pelos seguintes critérios:

- terem mostrado resultados satisfatórios no registro de imagens de satélite e aeronaves;
- apresentarem bom potencial no registro de imagens multi-temporais e de diferentes sensores, e
- terem apresentado uma avaliação dos experimentos.

O estudo comparativo é apresentado nas Tabelas 2.1 e 2.2. O símbolo representado por "—" nas tabelas significa que não foi possível obter tal ítem de informação no artigo pesquisado.

De acordo com Fonseca e Manjunath (1996) os casos de registro mais difíceis são aqueles em que as imagens

- 1) possuem resoluções espaciais diferentes;
- 2) são adquiridas em períodos de tempo diferentes, apresentando mudanças que afetam o processo de casamento;
- são afetadas por ruído, como o caso do ruído speckle nas imagens de radar, que dificulta a extração de informação;
- 4) são obtidas em diferentes faixas do espectro ou sob condições de iluminação diferentes e portanto, o conteúdo de informação de um mesmo objeto de cena aparece diferentemente nas imagens.

Todas estas situações são muitos comuns em aplicações na área de sensoriamento remoto. Com relação ao problema da resolução diferente nas imagens, tem-se sugerido que estas sejam corrigidas geometricamente e reamostradas para um mesmo tamanho de pixel para reduzir os erros de registro. Uma outra alternativa é usar atributos invariantes à escala no processo de casamento. Mas alguns experimentos têm mostrado que, para algumas situações, grandes diferenças de escala entre as imagens inviabilizam esta técnica.

Para reduzir o problema do ruído *speckle* nas imagens de radar, dois métodos têm sido sugeridos: (1) reduzir o efeito do ruído *speckle* antes da fase de extração de feições, e (2) usar a técnica de multiresolução durante o processo de extração de feições ou no processo de casamento destas. Filtros morfológicos, geométricos e estatísticos têm sido usados para reduzir o efeito do ruído *speckle* (Crimmins, 1985; Safa e Flouzat, 1989; Busse, 1995; Sant'Anna, 1995, Frery et al., 1997).

O registro de imagens multi-temporais é uma tarefa difícil por causa das mudanças temporais, justamente aquelas nas quais se tem interesse. Além disso, as mudanças da elevação solar produzem padrões de sombra que podem levar ao registro incorreto se os pontos de controle forem escolhidos nestas regiões.

As diversas metodologias de registro têm usado diferentes técnicas de realce de bordas, segmentação, casamento de feições e tipos de funções de transformação. A escolha de cada uma destas técnicas depende, em grande parte, da aplicação, das características dos sensores e da natureza da região imageada na superfície da Terra. Esta pluralidade de situações faz com que um único algoritmo não seja capaz de funcionar satisfatoriamente para todos os tipos de dados e aplicações.

# **TABELA 2.1 -**COMPARAÇÃO DOS MÉTODOS DE REGISTRO — PARTE I

Referência	Sensores	Feições	Processo de
			casamento
Goshtasby et al. (1986)	HCMM, TMS	contornos	método de
	Landsat2- MSS	fechados	clustering
Ton/Jain (1989)	Landsat- TM	regiões	relaxação
Ventura et al. (1990)	Landsat5- TM	regiões	casamento
	Spot, mapa		estrutural,
			árvore lógica
Wu/Maitre (1990)	Seasat SAR	pontos de	multiresolução,
	Spot	inflexão nos contornos	casamento de pontos
Caves et al. (1991)	mapa, SAR	conjunto de	correlação
		linhas	
Chandra (1992)	sensor ótico	coeficientes da TW	correlação
Cideciyan et al. (1992)	câmera	baseado em	correlação
		área	
Flusser/Suk (1994)	Spot, TM	contornos	atributos de
		fechados	forma, regra VMP
Toth/Schenk (1992)	Landsat- TM	mudanças na	atributos de forma
	mapa	curvatura	
Kher/Mitra (1993)	SAR	baseado em	Transf. Fourier
		área	
Dana e Anadan (1993)	sensores óticos	contornos	modelo elástico
			(gradiente)
Djamdji et al. (1993)	SPOT,Landsat-TM	coeficientes da	casamento iterativo
	Landsat-MSS	transf. wavelet	
Zheng/Chellapa (1993)	sensor ótico	transf. Gabor	correlação
	câmera		multiresolução
Moigne (1994)	Landsat-TM,	coeficientes da	correlação
	visual camera	transf. wavelet	multiresolução
Li et al. (1995a)	Landsat- TM	contornos fechados	correlação,
	Seasat, Spot	e abertos	atributos de forma
Corvi and Nicchiotti (1995)	SAR	coeficientes da	multiresolução
	sensor ótico	transf. wavelet	
Li e Zhou (1996)	images aéreas	pontos de borda	correlação
Deubler e Olivo (1997)	imagens médicas	coeficientes da	casamento por blocos
		transf. wavelet	

## TABELA 2.2 -COMPARAÇÃO DOS MÉTODOS DE REGISTRO — PARTE II

Referência	Número de exemplos	Transformação	Méritos e Limitações
Goshtasby et al. (1986)	2	afim polinomial	- imagens devem conter objetos bem definidos.
Ton/Jain (1989)	3	(RT)	- imagens com objetos bem definidos são usados.
			- pode ser adaptado para uma dada aplicação
Ventura et al. (1990)	1	polinomial	- flexibilidade na escolha dos atributos;
,			- a escolha dos atributos e regras lógicas;
			- deve ser auxiliada por especialista.
Wu/Maitre (1990)	1	(RTS)	- feições especiais são usadas: linhas costeiras.
Caves et al. (1991)	1	_	- não mostra ou avalia os resultados com detalhes;
Chandra (1992)	1	_	- desempenho comparável ao método de
			correlação mesmo na presença de ruído.
Cideciyan et al. (1992)	6	(RTS)	- bom desempenho em imagens ruidosas.
Flusser/Suk (1994)	1	afim genérica	- imagens devem conter objetos bem definidos.
Toth/Schenk (1992)	3	_	- feições específicas são usadas:regiões de água,
			culturas agrícolas e ruas.
Kher/Mitra (1993)	1	afim (rotação,	- tolerante ao ruído
		translação	- eficiente (tempo).
Dana/Anandam (1993)	3	afim	- adequado para imagens com pequeno deslocamento.
			- requer muito espaço em disco;
Djamdji et al. (1993)	5	polinômio	- imagens de sensores diferentes;
			- resultados muito bons.
Zheng/Chellapa (1993)	8	(RTS)	- tolera grande rotação e translação.
Moigne (1994)	2	(R)	- imagens somente com diferença de rotação.
			- imagens de sensores diferentes.
			- contornos bem definidos devem ser detectados.
Li et al. (1995)	6	(RTS)	- deslocamento entre as imagens de radar e
			óticas deve ser menor do que 5 pixels.
			- para pequenas distorções tolera grande rotação
Corvi/Nicchiotti (1995)	_	afim	e translação
			- tolera uma diferença de 10% na escala.
			- tolera pequena rotação e escala
Li e Zhou (1996)	3	polinômio	- computacionalmente eficiente
			- requer muitos pâmetros de ajuste
			- a forma dos objetos deve ser globalmente
			conservada
			- rotação no máximo de 20 graus
Deubler e Olivo (1997)	2	(RT)	- objeto de interesse deve ser centrado na imagem
			- condições de aquisição (iluminação) não deve mudar
			de imagem para imagem
			- tolera deformação global e local nas imagens.

## CAPÍTULO 3

#### **FUNDAMENTOS TEÓRICOS**

Como a transformada wavelet constitui uma das principais ferramentas deste trabalho, apresentamos neste Capítulo um resumo da sua teoria. O resumo é baseado principalmente nos trabalhos de Mallat (1989a,b), Mallat e Zhong (1992b) e Mallat e Hwang (1992). A teoria da transformada wavelet tem um formalismo matemático complexo e, portanto, procuramos apresentá-la sob o ponto de vista da teoria de filtros digitais e do conceito de multi-resolução para facilitar a sua compreensão. A teoria de wavelet sob o ponto de vista de processamento de sinais pode ser vista em Mallat (1998).

O objetivo é mostrar como a transformada wavelet bidimensional é computacionalmente implementada usando uma estrutura de pirâmide similar ao esquema de banco de filtros espelhados em quadratura (QMF) usado em processamento digital de sinais (Fliege, 1994; Vaidyanathan, 1993). Para uma melhor compreensão do esquema de banco de filtros, uma breve introdução à análise em multi-resolução é apresentada a seguir. Os apêndices apresentam algumas definições matemáticas básicas que são úteis para a compreensão deste Capítulo.

#### 3.1 Análise em Multi-resolução

Seja um sinal f e sua transformada de Fourier F, cuja magnitude é conhecida como o espectro do sinal, ilustrados na Figura 3.1. A noção de resolução de um sinal está relacionada com o conteúdo em freqüência do sinal da seguinte forma. Quanto maior a largura da banda de freqüência 2W do sinal (Figura 3.1b), mais componentes em altas freqüência estão presentes neste sinal (mais detalhes), e portanto melhor a resolução do sinal. Por outro lado, quanto menor a banda de freqüência de um sinal, menor o contéudo de informação e portanto menos detalhes o sinal possui. Dizemos que quando o contéudo de informação de um sinal diminui, a resolução deste sinal também diminui.



Fig. 3.1 - A função  $A\frac{\sin(\pi 2Wt)}{\pi 2Wt}$  e seu espectro de frequência limitado. FONTE: adaptada de Carlson (1981, p. 42).

Supondo que a função h seja a resposta impulsiva de um filtro, a filtragem linear de um sinal f por este filtro pode ser representada pela operação de convolução, resultando no sinal g:

$$g(x) = f(x) * h(x) = \int_{-\infty}^{+\infty} f(\xi)h(x - \xi)d\xi.$$
 (3.1)

A transformada de Fourier da Equação 3.1 é

$$G(w) = F(w).H(w) , \qquad (3.2)$$

onde H é a resposta em freqüência do filtro, e F e G são as transformadas de Fourier de f e g, respectivamente. Dependendo da característica da resposta em freqüência do filtro H, o filtro recebe diferentes denominações: passa-baixas, passa-altas, passa-banda e rejeita-faixas. O nosso interesse aqui se restringe aos filtros passa-baixas e passa-altas. Os filtros passa-baixas eliminam as componentes de alta freqüência do sinal e, portanto, diminui a sua resolução mantendo somente as componentes de baixas freqüências. Por outro lado, os filtros passa-altas deixam passar somente as componentes de altas freqüências do sinal, mantendo a informação de detalhes do sinal e eliminando a parte suave do sinal.

Dado um sinal original discreto  $f(n), n \in \mathbb{Z}$ , pode-se gerar um novo sinal y de baixa resolução filtrando este sinal com um filtro passa-baixas  $h_0$  e sub-amostrando o a uma taxa k. O processo de sub-amostragem na taxa k significa que a cada k

amostras uma é mantida. Assim, para uma taxa de sub-amostragem k=2 o sinal resultante terá a metade das amostras do sinal original. Então, o sinal resultante y pode ser representado por:

$$y(n) = \sum_{i=-\infty}^{+\infty} f(2n-i)h_0(i) . \tag{3.3}$$

A mudança de resolução do sinal é obtida pela filtragem passa-baixas e a de escala pela sub-amostragem na taxa 2.

Agora, a partir do sinal y, pode-se obter uma aproximação a do sinal original f. Isto pode ser obtido pelas seguintes operações. Primeiro inserem-se zeros entre cada duas amostras, ou seja, dobra-se a taxa de amostragem do sinal para que este tenha o mesmo número de amostras do sinal original:

$$y'(2n) = y(n), \quad y'(2n+1) = 0.$$
 (3.4)

A seguir o sinal y' é interpolado com um filtro de resposta impulsiva  $g_0$  para obter o sinal aproximado

$$a(n) = \sum_{k=-\infty}^{+\infty} y'(n-k)g_0(k) . {(3.5)}$$

Normalmente os sinais f e a não são iguais. A diferença entre estes sinais

$$d(n) = f(n) - a(n), \tag{3.6}$$

representa a informação de detalhe perdida no processo de filtragem. A Figura 3.2 mostra este esquema de filtragem.

A separação de um sinal original f em um sinal de aproximação grosseira a mais a informação de detalhe d é conceitualmente muito importante. Este procedimento está relacionado com a análise de sinais em multi-resolução que, quando repetido de modo iterativo, cria uma hierarquia de sinais de baixa resolução e em diferentes escalas. Por causa desta hierarquia e do fato que os sinais sub-amostrados tornam-se

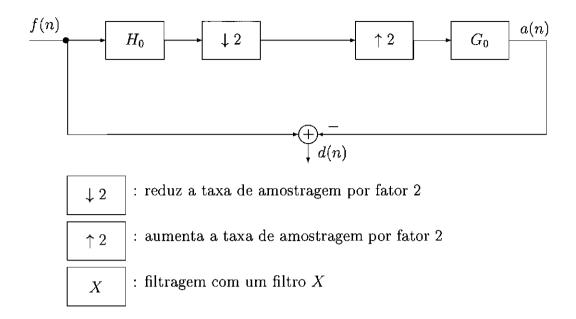
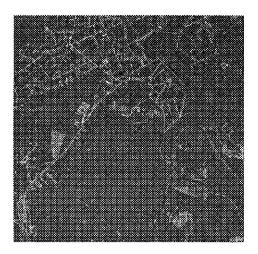


Fig. 3.2 - Esquema de pirâmide. Decomposição de um sinal f em dois sinais: aproximado a e detalhe d. O sinal original é igual a adição dos sinais a e d.

cada vez mais curtos (ou imagens menores), tais esquemas são chamados de pirâmide de sinais ou imagens. A Figura 3.3 ilustra uma pirâmide de imagens em três níveis de resolução. Esta sequência de imagens com resolução e escala diferentes foi obtida através das operações de filtragem e sub-amostragem em cascata na taxa 2. O filtro passa-baixas usado, neste exemplo, é um filtro 2-D separável nas linhas e colunas, com os coeficientes iguais a  $\{0,0625,0,25,0,375,0,25,0,0625\}$ .

Este conceito de pirâmide de multi-resolução pode ser generalizado usando o esquema de codificação sub-banda usado em processamento digital de sinais (Fliege, 1994). Neste esquema, dois operadores são usados para processar o sinal. O primeiro é um filtro passa-baixas (suavização)  $h_0$  e o outro é um filtro passa-altas (realce de detalhes)  $h_1$ .

Assim, um banco de filtros de análise  $h_0$  e  $h_1$  decompõe o sinal de entrada f em dois novos sinais  $f_0$  e  $f_1$ , respectivamente. Os sinais  $f_0$  e  $f_1$  são, respectivamente, os sinais de aproximação e detalhe discutidos anteriormente. A diferença está no modo



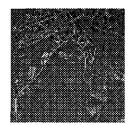






Fig. 3.3 - Pirâmide de imagens (3 níveis de resolução): Da esquerda para direita a resolução da imagem diminui por um fator de 2: (a) imagem original, 512x512 pixels; (b) imagem no nível 1, 256x256 pixels; (c) imagem no nível 2, 128x128 pixels e (d) imagem no nível 3, 64x64 pixels. Estas imagens são obtidas pelos processos de filtragem passa-baixas e sub-amostragem em cascata.

de geração do sinal detalhe que, ao invés de calcular o sinal diferença, usa um filtro passa-altas. Os sinais resultantes  $f_0$  e  $f_1$  são processados por um banco de filtros de síntese  $g_0$  e  $g_1$ , que reconstrói o sinal de saída  $\tilde{f}$ . O esquema computacional para os processos de decomposição e reconstrução do sinal usando os bancos de filtros de análise e síntese é ilustrado na Figura 3.4. Este esquema é conhecido por codificação sub-banda (SBC).

Os bancos de filtros são projetados para produzir um sinal de saída tão próximo quanto possível, ou mesmo exatatamente igual, ao sinal de entrada. No caso do sinal reconstruído ser exatamente igual ao sinal original, os bancos de filtros devem atender a algumas restrições específicas. Neste caso, os filtros são ditos terem propriedades de reconstrução perfeita.

A análise em multi-resolução tem sido muito utilizada em visão por computador e em

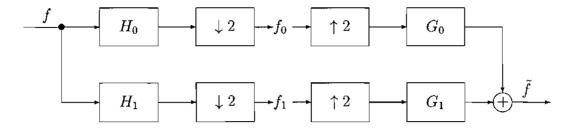


Fig. 3.4 - Esquema de Codificação Sub-banda (SBC). As aproximações do sinal correspondentes as componentes de baixas e altas freqüências são calculadas pelos filtros de análise  $H_0$  e  $H_1$ , respectivamente. Os filtros de síntese  $G_0$  e  $G_1$  reconstroem o sinal original.

processamento digital de imagens. Isto porque geralmente, os objetos nas imagens que se quer reconhecer têm diferentes tamanhos e é difícil definir a priori qual é a resolução mais adequada para analisar uma imagem. Algumas técnicas têm sido desenvolvidas para calcular a decomposição de um sinal em diferentes resoluções (Burt e Adelson, 1983; Mallat, 1989a; Toet, 1989). Em particular, a transformada wavelet tem sido muito usada em várias aplicações de análise em multi-resolução (Mallat e Zhong, 1992a; Chang e Kuo, 1993; Zhou e Ghosh, 1994).

A próxima seção descreve, de forma suscinta, a transformada wavelet, sua relação com a análise em multi-resolução e como ela pode ser implementada a partir de filtros espelhadas em quadratura (QMF) (Vaidyanathan, 1993; Fliege, 1994).

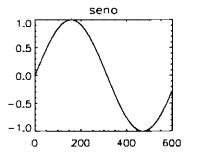
## 3.2 Wavelets

As wavelets têm sido um assunto bastante estudado e uma técnica usada em várias aplicações. A teoria de wavelets foi primeiramente usada para representar sinais sísmicos em termos de translações e dilatações de uma função fixa (Strang, 1989). Ela se mostrou mais eficiente do que a tradicional transformada de Fourier, apesar de não ser tão bem compreendida na época (Morlet, 1982).

As wavelets são funções que satisfazem certos requerimentos matemáticos e são usa-

das para representar funções, especialmente funções que são localizadas no tempo e na freqüência. O nome wavelet vem do requerimento de que estas funções deveriam ter integral zero, oscilando (fazendo ondinhas) acima e abaixo do eixo x. A conotação diminutiva de wavelet (que veio da palavra em francês ondelette sugere que a função deva ser bem localizada no espaço, sendo capaz assim de representar pequenas variações de um sinal. Os outros requerimentos são técnicos e necessários para assegurar os cálculos da transformada wavelet direta e inversa de forma rápida e simples.

Similares aos senos e cossenos na análise de Fourier, as wavelets são usadas como funções bases na representação de outras funções. O sinal é decomposto em versões dilatadas e deslocadas no tempo (espaço) da wavelet básica. Enquanto as senóides não tem duração finita, são suaves e tem comportamento previsível, as wavelets têm duração finita e tendem a ser irregulares e assimétricas, como ilustra a Figura 3.5.



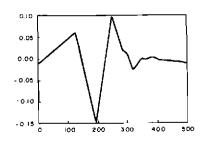


Fig. 3.5 - Funções seno e wavelet de Daubechies.

As principais diferenças entre a análise de Fourier e wavelets são (Strang, 1993):

• As wavelets permitem realizar uma análise localizada de um sinal. Por exemplo, os coeficientes wavelets são capazes de mostrar a exata posição de uma descontinuidade do sinal, o que não é possível através da análise de Fourier. Além disso, as componentes são obtidas através da filtragem de uma banda de freqüência do sinal original. Por isso dizemos que a wavelet tem boa localização nos domínios do espaço e da freqüência (Rioul e Vetterli, 1991).

- A transformada wavelet não possui um único conjunto de funções bases como a transformada de Fourier, que utiliza somente as funções senos e cossenos.
   As funções bases são diferentes na localização e na escala.
- Muitas classes de funções podem ser representadas através de wavelets de uma maneira compacta. Por exemplo, funções com descontinuidades e funções com grandes picos (imagens com bordas agudas e outras descontinuidades) geralmente podem ser representadas com bem menos funções bases wavelets do que as funções senos e cossenos para conseguir uma aproximação comparável.

#### 3.2.1 Transformada Wavelet

Seja  $\psi \in L^2(\mathbb{R})$  uma função complexa. A wavelet básica  $\psi$  satisfaz,

$$\int_{-\infty}^{+\infty} \psi(x)dx = 0. \tag{3.7}$$

Seja  $\Psi$  a transformada de Fourier de  $\psi$ . A condição acima implica que  $\Psi(0)=0$ . Então a função  $\psi$  pode ser interpretada como a resposta impulsiva de um filtro passa-banda.

A família de funções

$$\psi_{a,b}(x) = |a|^{-1/2} \psi(\frac{x-b}{a}) \ com \ a, b \in \mathbb{R}, a \neq 0$$
(3.8)

gerada a partir da wavelet básica dilatada de a e transladada de b são chamadas de wavelets. O índice de escala a está relacionada com a largura da função wavelet e o índice b com a sua posição no domínio espacial.

Seja um sinal mensurável e de energia finita,  $f \in L^2(\mathbb{R})$ . A transformada wavelet contínua de f é definida pelo produto de convolução

$$W_{a}f(x) = |a|^{-1/2} \int_{-\infty}^{+\infty} f(u)\psi(\frac{u-x}{a})du$$
  
=  $f(x) * \psi_{a}(-x)$ , (3.9)

onde  $\psi_a(x) = |a|^{-1/2} \psi(\frac{x}{a})$ . Esta operação produz os coeficientes wavelet  $W_a f$  na escala a e na posição x.

A wavelet  $\psi_{a,b}$  pode ser interpretada como uma resposta impulsiva de um filtro passa-banda, e a transformada wavelet de um sinal em uma dada escala como uma convolução deste sinal com o filtro dilatado. A translação corresponde ao deslocamento do filtro sobre o sinal, e o escalonamento, que é freqüentemente chamado de dilatação no contexto de wavelets, corresponde à mudança de largura da banda de freqüência do filtro. Desta forma, o processamento pode ser realizado em diferentes escalas ou resoluções, e em diferentes regiões do sinal. Assim, a transformada wavelet é uma ferramenta que decompõe dados ou funções em diferentes componentes de freqüência e então estuda cada componente com uma resolução condizente com a sua escala.

Para ilustrar esta operação, vamos reproduzir alguns passos de como se obtém os coeficientes de *wavelet* (Misiti et al, 1996), ilustrados na Figura 3.6:

- Escolhe-se uma wavelet e faz-se sua comparação com uma porção inicial do sinal original (Figura 3.6a);
- 2) Calcula-se um número C, que representa o grau de correlação entre a parte do sinal analisado e a wavelet. Quanto maior este número maior é o grau de similaridade entre a parte do sinal analisada e a wavelet (Figura 3.6a);
- 3) Desloca-se a wavelet para direita e repetem-se os passos 1 e 2 até que todo o sinal seja coberto (Figura 3.6b);
- 4) Aumenta-se a escala da função wavelet e repetem-se os passos de 1 a 3 (Figura 3.6c);
- 5) Repetem-se os passos 1-4 para todas as escalas.

No final deste processo tem-se um conjunto de coeficientes produzidos em diferentes escalas e porções do sinal. Quanto maior a escala, maior a porção do sinal que está sendo analisada, e portanto características mais grosseiras (componentes de baixa

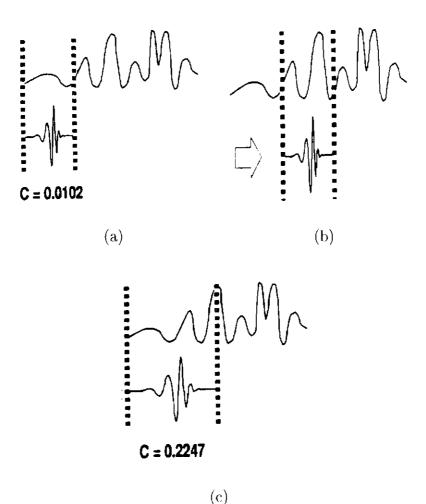


Fig. 3.6 - Cálculo da transformada wavelet. (a) sinal a ser decomposto e a wavelet;
(b) a wavelet se desloca para a direita e (c) a wavelet é escalonada.
FONTE: adaptada de Misiti et al (1996, p. 1.11).

freqüência) do sinal são medidas pelos coeficientes. Similarmente, detalhes mais finos (componentes de alta freqüência) são analisados quando a escala da wavelet é menor. É importante observar que os valores dos coeficientes dependem do formato da wavelet, e portanto a sua escolha deve ser feita com um certo critério, de acordo com o problema a ser resolvido.

No caso da transformada wavelet contínua, os parâmetros de dilatação e translação variam de forma contínua. Em aplicações práticas, os parâmetros a e b devem ser discretizados. A escolha mais comum é usar uma grade diádica, ou seja  $a=2^j$  e  $b=k.2^j$  com  $j,k\in \mathbb{Z}$ , sem modificar muito as propriedades da transformada (Mallat

e Zong, 1992b). Em geral, quanto menos valores para a e b se deseja usar, mais restritiva tornam-se as condições sobre as wavelets. A escolha de valores especiais de a e b naturalmente conecta a análise em multi-resolução, na área de processamento digital de sinais, com o mundo das wavelets.

A transformada wavelet de f na escala  $a = 2^j$  ć dada por:

$$W_{2^{j}} f(x) = f(x) * \psi_{2^{j}}(-x)$$

$$= \int_{-\infty}^{+\infty} f(u) \psi_{2^{j}}(u-x) du , \qquad (3.10)$$

onde  $\psi_{2^j}(x) = 2^{-j/2}\psi(2^{-j}x)$ . Pode-se mostrar que para uma classe particular de wavelets  $\psi(x)$ , a família de funções  $[\psi_{2^j}(x-2^{-j}n)]_{(n,j)\in\mathbb{Z}^2}$  é uma base ortonormal completa de  $L^2(\mathbb{R})$ . Consequentemente, as amostras fornecem uma representação completa da função f (Mallat e Hwang, 1992). Além disso, as amostras no espaço da transformada wavelet tendem a apresentar uma correlação reduzida. A propriedade de ortogonalidade garante que a similaridade da informação de detalbes, obtida nos diferentes níveis, seja causada pelas propriedades da imagem e não pela redundância da transformada aplicada como é o caso de outros métodos de multi-resolução, tais como a pirâmide laplaciana (Burt e Adelson, 1983).

A sequência de funções

$$Wf = [W_{2^j} f(x)]_{j \in \mathbb{Z}}$$
(3.11)

é conhecida por transformada wavelet diádica.

Na escala  $2^j$ , a função  $W_{2^j}f$  é contínua e sua transformada de Fourier é dada por:

$$\Omega_{2i}(w) = F(w) \cdot \Psi(2^j w) , \qquad (3.12)$$

onde  $\Omega_{2^j}$ , F e  $\Psi$  são as transformadas de Fourier de  $W_{2^j}f$ , f e  $\psi$ , respectivamente.

Se estabelecermos que existem duas constantes positivas A e B tal que

$$\forall w \in \mathbb{R}, A \le \sum_{j=-\infty}^{+\infty} |\Psi(2^j w)|^2 \le B, \tag{3.13}$$

pode-se assegurar que todo o eixo de freqüências é coberto por dilatações de  $\psi$  nas escalas  $(2^j)_{j\in \mathbb{Z}}$  (Mallat e Zhong, 1992b). Qualquer wavelet satisfazendo esta condição é chamada de wavelet diádica. Se a condição 3.13 é válida, então a função f pode ser recuperada a partir de sua transformada wavelet diádica. A wavelet de reconstrução  $\xi$  é uma função cuja transformada de Fourier satisfaz a condição:

$$\sum_{j=-\infty}^{+\infty} \Psi(2^j w) \chi(2^j w) = 1, \tag{3.14}$$

onde  $\chi$  é a transformada de Fourier de  $\xi$ . Assim, a função f pode ser recuperada usando-se

$$f(x) = \sum_{j=-\infty}^{+\infty} W_{2j} f(x) * \xi_{2j}(x). \tag{3.15}$$

### 3.2.2 Transformada Wavelet Diádica Discreta

Na prática não se pode calcular a transformada wavelet em todas as escalas  $2^j$  para j variando de  $-\infty$  a  $+\infty$ . Para se obter uma transformada discreta devemos reconhecer que as escalas têm um limite inferior e superior para um sinal discreto. Vamos supor que a escala mais fina seja igual a 1 e que  $2^L$  seja a escala maior.

Seja  $\phi$  uma função de suavização tal que sua transformada de Fourier satisfaça:

$$|\Phi(w)|^2 = \sum_{j=1}^{+\infty} \Psi(2^j w) \chi(2^j w).$$
 (3.16)

Supõe-se que  $\chi$  seja uma função tal que  $\Psi\chi$  seja uma função real, positiva. Das Equações 3.14 e 3.16, pode-se provar que (Apêndice B)

$$\lim_{w \to 0} \Phi(w) = 1 \ . \tag{3.17}$$

Deste resultado pode-se concluir que a energia de  $\Phi$  está concentrada nas baixas

freqüências e que, portanto,  $\phi$  é uma função de suavização (filtro passa-baixas).

Seja  $A_{2^j}$  o operador de suavização que calcula a aproximação  $A_{2^j}f$  de um sinal f na resolução  $2^j$ :

$$A_{2j} f(x) = f(x) * \phi_{2j}(x) = \int_{-\infty}^{+\infty} f(u)\phi_{2j}(x-u)du ,$$
 (3.18)

onde,  $\phi_{2^j}(x) = 2^{-j/2}\phi(2^{-j}x)$ .

Como  $\phi$  é a resposta de um filtro passa-baixas, detalhes presentes em f são removidos por este filtro gerando um outro sinal com resolução mais baixa. Por isso  $A_{2^j}f$  é chamada de aproximação de f na resolução  $2^j$ . Quando processamos o sinal em escalas maiores  $(j \to L)$ , maior conteúdo de informação é perdido e a resolução do sinal aproximado diminui. De modo inverso, quando diminuimos o fator de escala  $(j \to 1)$  o sinal aproximado converge para o sinal original.

Calculando a transformada de Fourier da Equação 3.18 para os níveis de resolução 1 e  $2^L$  temos que

$$S_1 f(w) = F(w) \cdot \Phi(w)$$

$$S_{2^L} f(w) = F(w) \cdot \Phi(2^L w) . \tag{3.19}$$

Da Equação (3.16), pode-se obter a seguinte expressão:

$$|\Phi(w)|^{2} = \sum_{j=1}^{L} \Psi(2^{j}w)\chi(2^{j}w) + \sum_{j=L+1}^{+\infty} \Psi(2^{j}w)\chi(2^{j}w)$$
$$= \sum_{j=1}^{L} \Psi(2^{j}w)\chi(2^{j}w) + |\Phi(2^{L}w)|^{2}.$$
(3.20)

A partir da Equação 3.20 e usando as Equações 3.12, 3.19 e o teorema de Parseval, obtém-se a seguinte equação de conservação de energia (Apêndice B):

$$||A_1 f(x)||^2 - ||A_2 f(x)||^2 = \sum_{j=1}^L W_{2j} f(x) * \xi_{2j}(x)$$
. (3.21)

Este resultado mostra que as componentes em alta freqüências de f removidas pela operação de suavização na escala  $2^L$  podem ser recuperadas pela transformada wavelet diádica  $[W_{2^j}f]_{1\leq j\leq L}$  entre as escalas  $2^1$  e  $2^L$ . Ou seja, a transformada wavelet diádica  $[W_{2^j}f]$  entre os níveis 1 e  $2^L$  fornece os detalhes presentes em  $A_1f$  mas que foram removidos em  $A_{2^L}f$  devido a operação de suavização. O sinal  $W_{2^j}f$ , que contém a diferença de informações entre os sinais aproximados  $A_{2^j}f$  e  $A_{2^{j+1}}f$  é chamado de sinal detalhe.

#### 3.2.3 Representação em Wavelets para Sinais 1D

Na prática, o sinal a ser processado é uma seqüência de valores discretos. Então, pode-se provar (Mallat, 1989a) que o sinal discreto  $A_1^d f$  medido na resolução 1 é representado pelo conjunto de sinais discretos

$$\left(A_{2^L}^d f, \ (D_{2^j} f)_{1 \le j \le L}\right),$$
 (3.22)

onde

$$A_{2^{j}}^{d}f = (f(u) * \phi_{2^{j}}(-u))|_{(2^{-j}n)_{n \in \mathbb{Z}}}$$
(3.23)

e

$$D_{2^{j}}f = (f(u) * \psi_{2^{j}}(-u))|_{(2^{-j}n)_{n \in \mathbb{Z}}}.$$
(3.24)

Os sinais  $A_{2^j}^d f$  e  $D_{2^j} f$  são resultados da convolução do sinal f com as funções  $\phi_{2^j}(-x)$  e  $\psi_{2^j}(-x)$ , respectivamente, avaliados nas posições  $2^{-j}n$ . Ou seja, estes sinais podem ser interpretados como uma filtragem passa-baixas  $(\phi_{2^j})$  ou passa-altas  $(\psi_{2^j})$  de f seguida por uma amostragem uniforme na taxa  $2^j$ . Esta seqüência de sinais discretos é chamada de representação ortogonal por wavelets. Ela consiste de um sinal aproximado na menor resolução  $A_{2^L}^d f$  e dos sinais detalhes  $D_{2^j} f$  nas resoluções  $2^j$  para  $1 \le j \le L$ . Este resultado é muito importante, e está diretamente relacionado com a análise em multi-resolução e com a teoria de banco de filtros, mencionadas na Seção 3.1.

Se o sinal original tem N amostras, então os sinais discretos  $D_{2^j}f$  e  $A_{2^j}^df$  têm cada um  $2^{-j}N$  amostras. Assim, a representação por wavelet

$$\left(A_{2^L}^d f, \ (D_{2^j} f)_{1 \le j \le L}\right) ,$$
 (3.25)

tem o mesmo número de amostras do que o sinal aproximado  $A_1^d f$ . Isto ocorre porque a representação é ortogonal.

Mallat (1989a) mostrou que os sinais aproximados  $A_{2^j}^d f$  e detalhes  $D_{2^j} f$  podem ser calculados de forma similar ao esquema de banco de filtros. Seja h a resposta impusiva do filtro passa-baixas

$$h(n) = \int_{-\infty}^{+\infty} \phi(x)\phi(2x-n)dx , \qquad (3.26)$$

e o filtro simétrico  $\tilde{h}(n) = h(-n)$ . O sinal de aproximação  $A_{2^j}^d f$  na resolução  $2^j$  pode ser calculado convoluindo  $A_{2^{j-1}}^d f$  com o filtro passa-baixas  $\tilde{h}$ , e sub-amostrando o sinal resultante. Todas os sinais de aproximação  $(A_{2^j}^d f)_{1 \leq j \leq L}$  podem ser calculados a partir de  $A_1^d f$  repetindo o mesmo procedimento, como ilustra o diagrama de blocos na Figura 3.7.

Os sinais  $D_{2^j}f$  podem ser calculados convoluindo  $A_{2^{j-1}}^df$  com o filtro discreto passaaltas  $\tilde{g}(n) = g(-n)$ , onde

$$g(n) = \int_{-\infty}^{+\infty} \phi(x)\psi(2x - n)dx , \qquad (3.27)$$

usando o mesmo procedimento usado no cálculo de  $A^d_{2^j}f$ .

No processo de reconstrução (Figura 3.8), o sinal  $A^d_{2^{j-1}}f$  pode ser reconstruído colocando-se zeros entre cada duas amostras de  $A^d_{2^j}f$  e  $D^f_{2^j}$ , e efetuando-se as convolução dos sinais resultantes com os filtros h e g, respectivamente. O sinal discreto na escala mais fina  $A^d_1f$  é reconstruído repetindo o mesmo procedimento para todos os níveis  $1 \leq j \leq L$ .

A resposta impulsiva do filtro G está relacionada com a resposta impulsiva do filtro

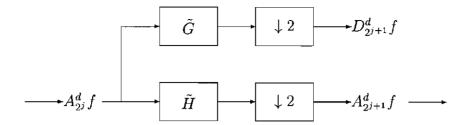


Fig. 3.7 - Decomposição de uma aproximação discreta  $A_{2^j}^d f$  em uma aproximação na resolução mais baixa  $A_{2^{j+1}}^d f$  e um sinal detalhe  $D_{2^{j+1}}^d f$ . FONTE: adaptada de Mallat (1989a).

H por (Mallat, 1989a):

$$g(n) = (-1)^{1-n}h(1-n). (3.28)$$

Em processamento digital de sinais G e H são conhecidos como filtros espelhados em quadratura (QMF), e o filtro G é o filtro espelho de H (Vaidyanathan, 1993).

## 3.2.4 Representação em Wavelets para Imagens

A representação por wavelet descrita na seção 3.2.3 pode ser facilmente estendida para o caso bidimensional. Considerando a função de suavização bidimensional  $\phi(x,y)$  separável nas direções x e y,  $\phi$  pode ser escrita como

$$\phi(x,y) = \phi(x)\phi(y) , \qquad (3.29)$$

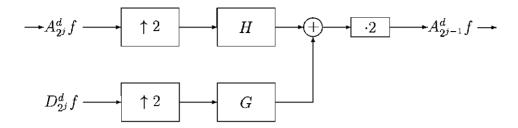
onde  $\phi(x)$  é a função de suavização unidimensional.

No caso bidimensional temos três wavelets para decompor o sinal:

$$\psi^{1}(x,y) = \phi(x)\psi(y)$$

$$\psi^{2}(x,y) = \psi(x)\phi(y)$$

$$\psi^{3}(x,y) = \psi(x)\psi(y).$$
(3.30)



 $\uparrow 2$  : coloca uma amostra entre cada duas amostras

: multiplica por 2

Fig. 3.8 - Reconstrução do sinal de aproximação discreta  $A^d_{2^{j-1}}f$  a partir de um sinal de aproximação na resolução mais baixa  $A^d_{2^j}f$  e do sinal detalhe  $D^d_{2^j}f$ .

FONTE: adaptada de Mallat (1989a).

Então, para uma imagem  $f \in L^2(\mathbb{R}^2)$ , as imagens de aproximação e detalhes na resolução  $2^j$  podem ser representados por:

$$A_{2^{j}}^{d} f = (f(x,y) * \phi_{2^{j}}(-x)\phi_{2^{j}}(-y))|_{(2^{-j}n,2^{-j}m)_{(n,m)\in\mathbb{Z}^{2}}}$$

$$D_{2^{j}}^{1} f = (f(x,y) * \phi_{2^{j}}(-x)\psi_{2^{j}}(-y))|_{(2^{-j}n,2^{-j}m)_{(n,m)\in\mathbb{Z}^{2}}}$$

$$D_{2^{j}}^{2} f = (f(x,y) * \psi_{2^{j}}(-x)\phi_{2^{j}}(-y))|_{(2^{-j}n,2^{-j}m)_{(n,m)\in\mathbb{Z}^{2}}}$$

$$D_{2^{j}}^{3} f = (f(x,y) * \psi_{2^{j}}(-x)\psi_{2^{j}}(-y))|_{(2^{-j}n,2^{-j}m)_{(n,m)\in\mathbb{Z}^{2}}}.$$
(3.31)

A decomposição wavelet de uma imagem fornece uma representação fácil de ser interpretada. A decomposição wavelet pode ser interpretada como a decomposição de um sinal em um conjunto de bandas de freqüências espaciais independentes. Os filtros usados na decomposição e síntese do sinal são os filtros  $\tilde{G}$ ,  $\tilde{H}$ , G e H, descritos na Seção 3.2.3. Fazendo uma análise no domínio da freqüência, a imagem  $A_{2^j}^d f$  corresponde as componentes de baixas freqüências da imagem original, e as imagens  $D_{2^j}^1 f$ ,  $D_{2^j}^2 f$  e  $D_{2^j}^3 f$  representam as componentes de altas freqüências ao longo das

direções vertical (bordas horizontais), horizontal (bordas verticais) e diagonal (cantos), respectivamente. As imagens  $A_{2^j}^d f$ ,  $D_{2^j}^1 f$ ,  $D_{2^j}^2 f$  e  $D_{2^j}^3 f$  são também chamadas, respectivamente, de bandas LL (Low-Low), LH (Low-High), HL (High-Low) e HH (High-High), de acordo com o tipo de filtragem usada. A Figura 3.9 mostra as disposições destas imagens na representação por wavelet em duas dimensões.

$\begin{vmatrix} A_{2^3}^d f & D_{2^3}^1 f \\ D_{2^3}^2 f & D_{2^3}^3 f \end{vmatrix}$	$D_{22}^{1}f$	$D^1_{2^1}f$
$D_{2^2}^2 f$	$D_{2^2}^3f$	221
$D^2_{2^1}f$		$D^3_{2^1}f$

Fig. 3.9 - Disposição das imagens  $D_{2^j}^k f(k=1,2,3)$  e  $A_{2^j}^d f$  na representação wavelet de uma imagem para três níveis de resolução (L=3).

A imagem original  $A_1f$  é computacionalmente representada por 3L+1 imagens

$$\left(A_{2^L}^d f, (D_{2^j}^1 f)_{1 \le j \le L}, (D_{2^j}^2 f)_{1 \le j \le L}, (D_{2^j}^3 f)_{1 \le j \le L}\right) . \tag{3.32}$$

Este conjunto de imagens é conhecido como uma representação por wavelets em duas dimensões. A imagem  $A^d_{2L}f$  é a aproximação grosseira da imagem original na resolução  $2^L$ , e as imagens  $D^k_{2j}f$  fornecem a informação dos detalhes da imagem original nos diferentes níveis de resolução. O número total de pixels nesta representação é igual ao número de pixels da imagem original. Para uma imagem original com N×N pixels, cada sub-imagem ou banda possui  $2^{-j}N\times 2^{-j}N$  pixels.

Similarmente ao caso unidimensional, a representação bidimensional pode ser com-

putada com um algoritmo de pirâmide, representado pelo diagrama de blocos na Figura 3.10. Em cada nível o sinal  $A^d_{2^j}f$  é decomposto em quatro novos sinais na resolução  $2^{j+1}$ ,  $A^d_{2^{j+1}}f$ ,  $D^1_{2^{j+1}}f$ ,  $D^2_{2^{j+1}}f$  e  $D^3_{2^{j+1}}f$ . O processamento é realizado nas linhas e colunas separadamente, usando os filtros  $\tilde{H}$  e  $\tilde{G}$  descritos na Seção 3.2.3. Primeiro, as linhas de  $A^d_2f$  são convoluídas com um filtro unidimensional. A cada duas colunas uma é mantida. O sinal resultante é convoluído com o outro filtro unidimensional na direção das colunas. Novamente, a cada duas linhas uma é mantida, e assim temos uma imagem com tamanho reduzido de 1/4. Este procedimento é repetido sucessivamente para os outros níveis até o último nível de resolução desejado.

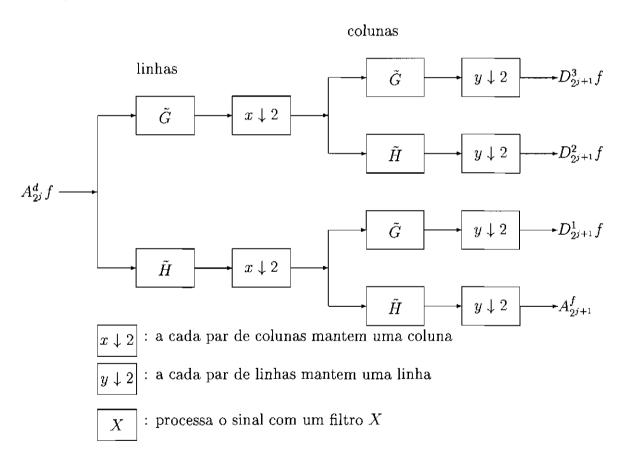


Fig. 3.10 - Decomposição de uma imagem  $A_{2^j}^df$  nas sub-imagens  $A_{2^{j+1}}^df$ ,  $D_{2^{j+1}}^1f$ ,  $D_{2^{j+1}}^2f$ ,  $D_{2^{j+1}}^2f$ . FONTE: adaptada de Mallat (1989a).

No processo de reconstrução a imagem é reconstruída a partir da representação por

wavelet com um algoritmo piramidal semelhante ao apresentado na Figura 3.11. Em cada nível  $2^j$ , a imagem  $A^d_{2^j}f$  é reconstruída a partir das sub-imagens  $A^d_{2^{j+1}}f$ ,  $D^1_{2^{j+1}}f$ ,  $D^2_{2^{j+1}}f$  e  $D^3_{2^{j+1}}f$ . Acrescentam-se colunas de zeros entre as colunas destas imagens, convoluem-se as linhas com um filtro unidimensional, adicionam-se linhas de zeros entre as linhas resultantes e convoluem-se as colunas com o outro filtro unidimensional. A imagem original  $A^d_1f$  é reconstruída pela repetição sucessiva deste procedimento para todos os níveis j=L,L-1,...,1. Se usarmos precisão de ponto flutuante para os sinais discretos na representação wavelet, a reconstrução é, em geral, de excelente qualidade.

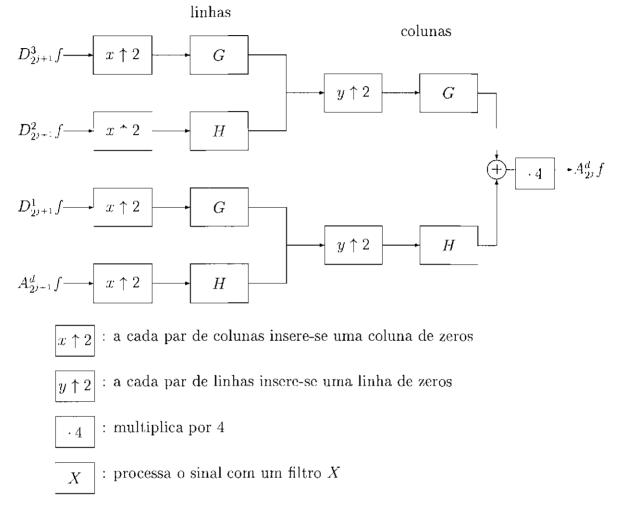


Fig. 3.11 - Reconstrução de uma imagem de aproximação discreta  $A^d_{2^j}f$  a partir de uma imagem de aproximação na resolução mais baixa  $A^d_{2^{j+1}}f$  e das imagens de detalhes  $D^k_{2^{j+1}}f(k=1,2,3)$ .

FONTE: Adaptada de Mallat (1989a).

# CAPÍTULO 4

# DETECÇÃO DE FEIÇÕES PONTUAIS

Na área de reconhecimento de padrões dois assuntos são muito importantes: a identificação e o casamento de feições. A Figura 4.1 mostra um esquema dos processos envolvidos na área de reconhecimento de padrões. A identificação envolve a detecção de feições úteis e a representação da informação sobre os objetos em uma estrutura de dados apropriada. O casamento consiste em encontrar uma representação, dentre todas as armazenadas, se existir, que mais se identifica com um dado padrão (Manjunath, 1991). Neste Capítulo, discutiremos somente o problema de identificação de feições.

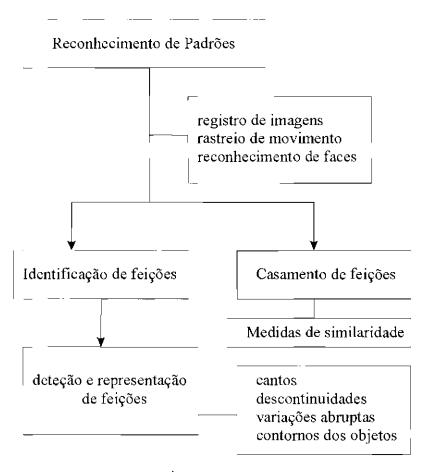


Fig. 4.1 - Esquema dos processos envolvidos na área de reconhecimento de padrões.

A detecção e representação de feições é um assunto básico na maioria dos problemas

de visão e processamento de imagens, tais como registro de imagens, rastreio de movimento (motion tracking), etc. Na literatura há muitos algoritmos que detectam feições salientes em imagens e usam as informações na vizinhança destas feições para várias aplicações (Kitchen e Rosenfeld, 1982; Manjunath e Chellapa, 1992; Chen et al, 1995; Lee et al., 1993; Manohar e Moigne, 1998). Feições típicas incluem cantos, pontos com descontinuidades significantes da curvatura e pontos com grandes variações de intensidade.

As feições de interesse em uma imagem, geralmente, estão presentes em diferentes tamanhos. Uma maneira eficiente para analisar tais feições é através da decomposição em multi-resolução da imagem, que fornece um esquema hierárquico elegante para a análise de imagens. O esquema de decomposição, conhecida como pirâmide Laplaciana, proposta por Burt e Adelson (1983) foi inicialmente usada para codificação de sinais. Mais recente é a aplicação da transformada wavelet para a análise de imagens em multi-resolução (Mallat, 1989a), que tem se mostrado muito eficiente para a análise local de sinais. A transformada wavelet tem sido usada em diferentes aplicações tais como classificação por textura (Chang e Kuo, 1993), segmentação de imagens (Zhou et al, 1994)), detecção de bordas (Mallat e Zhong, 1992a), codificação de imagens (Antonini et al., 1992), análise de imagens SAR (Teti e Kritikos, 1992)), filtragem de ruído (Xu et al., 1994) e reconhecimento de objetos (Sheng et al., 1993).

Existe uma importante conexão entre wavelets e detecção de bordas considerando que a transformada wavelet é bem adequada para a localização de variações abruptas do sinal. A relação entre a transformada wavelet e o problema de detecção de bordas, formalizada por Mallat e Zhong (1992a), é apresentada na Seção 4.1.

Na Seção 4.2 apresentamos um procedimento automático para extrair feições pontuais usando a transformada wavelet em diferentes níveis de resolução. Esta técnica é bastante simples, eficiente e depende basicamente de um parâmetro que controla o número de feições pontuais extraídas.

Na Seção 4.4 alguns resultados experimentais são apresentados para ilustrar o desempenho do método na extração de feições pontuais.

# 4.1 Detecção de Bordas através do Módulo dos Máximos da Transformada Wavelet

Seja uma função de suavização  $\theta$  a resposta impulsiva de um filtro passa-baixa 2-D, tal que  $\int \int \theta(x,y) dx dy \neq 0$ . Considere que esta função seja diferenciável. A primeira derivada de  $\theta$  pode ser decomposta em duas componentes  $\psi^1$  e  $\psi^2$  ao longo dos eixos x e y, respectivamente, dadas por

$$\psi^{1}(x,y) = \frac{\partial \theta(x,y)}{\partial x}$$

$$\psi^{2}(x,y) = \frac{\partial \theta(x,y)}{\partial y},$$
(4.1)

Por definição estas funções podem ser usadas como wavelets porque suas integrais são iguais a zero (Mallat e Zhong, 1992a).

Para uma função  $f \in L^2(\mathbb{R})$ , a transformada wavelet na escala  $a = 2^j$  definida com relação a estas duas wavelets tem duas componentes (Mallat e Hwang, 1992):

$$W_{2j}^{1} f(x,y) = f(x,y) * \psi_{2j}^{1}(x,y)$$

$$= f * (2^{j} \frac{\partial}{\partial x} \theta_{2j}(x,y))$$

$$= 2^{j} \frac{\partial}{\partial x} (f * \theta_{2j})(x,y)$$

$$W_{2j}^{2} f(x,y) = f(x,y) * \psi_{2j}^{2}(x,y)$$

$$= f(x,y) * (2^{j} \frac{\partial}{\partial y} \theta_{2j}(x,y))$$

$$= 2^{j} \frac{\partial}{\partial y} (f * \theta_{2j})(x,y) . \tag{4.2}$$

A Equação 4.2 pode ser reescrita na forma matricial

$$\begin{pmatrix} W_{2j}^{1}f(x,y) \\ W_{2j}^{2}f(x,y) \end{pmatrix} = 2^{j} \begin{pmatrix} \frac{\partial}{\partial x}(f*\theta_{2j})(x,y) \\ \frac{\partial}{\partial y}(f*\theta_{2j})(x,y) \end{pmatrix}$$
$$= 2^{j} \vec{\nabla}(f*\theta_{2j}(x,y)). \tag{4.3}$$

Portanto, as componentes da transformada  $wavelet W_{2^j}^1 f$  e  $W_{2^j}^2 f$  são proporcionais às duas componentes do vetor gradiente da imagem f suavizada por  $\theta_{2^j}$ , e caracterizam as singularidades ao longo das direções x e y, respectivamente (Mallat e Hwang, 1992). Podemos observar ainda que, quando a escala  $a = 2^j$  é grande, a convolução com  $\theta_{2^j}$  remove as pequenas variações do sinal. Isto implica que somente as grandes variações são detectadas.

O módulo do vetor gradiente, na escala  $2^{j}$ , é proporcional a

$$M_{2^{j}}(x,y) = \sqrt{\left|W_{2^{j}}^{1}f(x,y)\right|^{2} + \left|W_{2^{j}}^{2}f(x,y)\right|^{2}},$$
(4.4)

que é o módulo da transformada wavelet na escala  $2^{j}$ .

O ângulo  $\alpha$  entre o gradiente  $\vec{\nabla}(f*\theta)$  e a direção horizontal é dado por

$$\alpha_{2^{j}}(x,y) = tan^{-1} \left( \frac{W_{2^{j}}^{2} f(x,y)}{W_{2^{j}}^{2} f(x,y)} \right) . \tag{4.5}$$

Este ângulo indica localmente a direção para onde o sinal tem maior variação e é perpendicular às direções das bordas.

Mallat e Zhong (1992a) mostraram que os valores máximos de  $M_{2^j}$  caracterizam pontos de variações agudas de f, e que portanto podem ser usados para detectar pontos de borda de f na escala  $2^j$ . Este procedimento é semelhante ao usado por Canny (1986). Para cada escala  $2^j$  os pontos (x,y) para os quais a magnitude do gradiente  $M_{2^j}$  é localmente máxima ao longo da direção dada por  $\alpha_{2^j}$  são detectados. Estes pontos são denominados de módulos máximos da transformada wavelet e são considerados como pontos de borda na imagem. Desta forma, cada módulo máximo é caracterizado pela sua posição (x,y), módulo  $M_{2^j}$  e ângulo  $\alpha_{2^j}$ .

Mallat e Zhong (1992a) mostraram que os módulos máximos nos diferentes níveis de resolução fornecem uma representação da imagem completa e estável, chamada de representação por máximos da wavelet, suficiente para a reconstrução da imagem original. Além disso, com esta representação é possível desenvolver algoritmos unicamente baseado em bordas para resolver diversos problemas na área de pro-

cessamento de imagens, tais como codificação, discriminação de textura, realce e reconhecimento de padrões (Mallat e Zhong, 1992a; Laine e Fan, 1993).

# 4.2 Detecção de Feições Pontuais

A fase mais crtica do processo de registro é a determinação dos pontos de controle. Se o modelo de distorção escolhido é adequado, a precisão do registro é garantida pela escolha adequada dos pontos de controle.

Nesta seção propomos um método simples para selecionar, de forma automática, feições significativas na imagem que podem ser usadas como pontos de controle no registro. Estas feições são extraídas através dos máximos locais do módulo da transformada wavelet, para os diferentes níveis de resolução.

Moigne (1994), Djamdji et al (1993) e Corvi e Nicchiotti (1995) também usam a transformada wavelet para extrair feições salientes nas imagens, as quais são usadas no processo de registro. Moigne (1994) usa uma porcentagem (10%) do número total de pontos nas bandas LH e HL da decomposição wavelet como os coeficientes máximos para representar as feições salientes nas imagens. Djamdji et al (1993) usam os pontos cujos coeficientes da transformada wavelet, obtidos pelo algoritmo à trous (Bijaoui e Giudicelli, 1991), são máximos. Corvi e Nicchiotti (1995) usam os pontos cujos coeficientes da transformada wavelet são máximos ou mínimos para representar as feições salientes nas imagens. O processo de casamento é realizado separadamente para as feições de máximos e mínimos. Zheng e Chellapa (1993) usam a decomposição da função de Gabor no processo de extração de feições, que é computacionalmente mais caro e complexo.

Diferentemente, usamos os máximos locais da imagem módulo da transformada wavelet na direção do gradiente para extrair as feições pontuais na imagem. As sub-imagens de altas freqüências da decomposição wavelet são usadas para estimar os gradientes nas direções horizontal e vertical. A combinação das informações de bordas presentes nas sub-imagens de altas freqüências em uma única imagem facilita e acelera os processos de identificação e casamento de feições. Como usamos a

imagem módulo, os coeficientes máximos e mínimos são detectados usando o mesmo procedimento. Para a decomposição wavelet, usamos um algoritmo similar ao descrito na Seção 3.2.4, que é mais eficiente, em termos de espaço em disco e tempo de computação, do que o algoritmo usado por Djamdji et al (1993). A redução de espaço em disco e memória são fatores importantes no processamento de imagens de sensoriamento remoto que, geralmente, são muito grandes.

O método de extração de feições é realizado nas seguintes etapas:

- 1) detecção dos pontos de bordas;
- 2) eliminação de feições não significativas, ou seja, pontos de bordas fracas;
- 3) eliminação de pontos localizados em áreas de baixo contraste;
- 4) supressão dos pontos de não-máximos locais.

Na primeira fase, a imagem é processada em L níveis da decomposição wavelet. Em cada nível de resolução  $2^j$  as imagens módulo da transformada wavelet  $M_{2^j}$  e ângulo  $\alpha_{2^j}$  são calculadas, usando as Equações 4.4 e 4.5. Então, um ponto  $P(x_o, y_o)$  é considerado ponto de borda se

$$M_{2j}(x_o, y_o) > M_{2j}(x, y),$$
 (4.6)

onde (x, y) pertence à vizinhança (dois pontos) mais próxima de  $P(x_o, y_o)$  na direção indicada pelo ângulo  $\alpha(x_o, y_o)$ .

Seja um parâmetro  $au_{2^j}$  definido por

$$\tau_{2^j} = \beta \cdot \sigma_{2^j} + \mu_{2^j}, \tag{4.7}$$

onde  $\beta$  é uma constante, e  $\sigma_{2^j}$  e  $\mu_{2^j}$  são o desvio padrão e a média da imagem módulo da transformada wavelet no nível de resolução  $2^j$ , respectivamente. Uma feição no ponto  $P(x_o, y_o)$  é considerada significativa se

$$M_{2j}(x_o, y_o) > \tau_{2j} . (4.8)$$

Esta condição elimina as bordas fracas, e garante uma consistência das bordas em todos os níveis de resolução. Ou seja, se uma borda é forte, ela deve aparecer, praticamente, em todos os níveis de resolução. O número de feições significativas é controlado pelo parâmetro  $\beta$ , Equação 4.7, e seu valor inicial é definido pelo usuário.

Seja uma janela  $w_c \mathbf{x} w_c$  (onde  $w_c$  é inteiro e ímpar) de pontos centrada no ponto  $P(x_o, y_o)$ ,  $\sigma_w$  o desvio padrão dos níveis de cinza da imagem dentro desta janela. Seja uma medida de contraste do ponto  $P(x_o, y_o)$  definida por (Gonzalez e Woods, 1992)

$$CO(x_o, y_o) = 1 - \frac{1}{1 + \sigma_w}.$$
 (4.9)

Uma feição no ponto  $P(x_o, y_o)$  é armazenada na imagem se e somente se:

$$CO(x_o, y_o) > T_w , \qquad (4.10)$$

onde  $T_w$  é um limiar acima do qual o contraste do ponto é considerado alto. Esta condição garante que os pontos escolhidos estejam localizados em áreas de alto contraste reduzindo, assim, o número de casamentos falsos no processo de casamento, como será visto no Capítulo 5.

Os máximos locais são calculados de tal forma que

$$M_{2^{j}}(x_{o}, y_{o}) = \max_{(x,y) \in V_{p}} \{ M_{2^{j}}(x,y) \}, \qquad (4.11)$$

onde  $V_p$  é a vizinhança de  $P(x_o, y_o)$ , definida como uma janela de pontos  $w_p \times w_p$ . Esta condição garante que o ponto de feição seja único na sua vizinhança.

#### 4.3 Detalhes de Implementação

A implementação de um algoritmo nem sempre é trivial. O programa que codifica a solução de um problema é o resultado de muitos experimentos e refinamentos para que o problema seja resolvido na prática. Na área de processamento digital de imagens, como os dados são digitais, muitas aproximações são realizadas. Nesta

seção, apresentamos alguns detalhes da implementação do algoritmo de detecção de feições pontuais apresentado na seção anterior. Por convenção, adotamos o sistema de coordenadas da Figura 4.2 para representar uma imagem digital f.

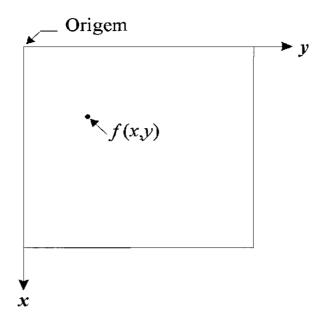


Fig. 4.2 - Convenção adotada para o sistema de coordenadas x e y na representação da imagem digital f.

Para a decomposição wavelet usamos uma wavelet do tipo spline cúbica, descrita em Mallat e Hwang (1992). A Figura 4.3 mostra os gráficos de uma wavelet spline cúbica e da função de suavização  $\phi$  correspondente a esta wavelet, que é também uma spline cúbica com um suporte de tamanho 4.

A implementação da transformada wavelet é realizada usando um algoritmo similar ao algoritmo apresentado na Seção 3.2.4, no qual, em cada nível de resolução quatro sub-imagens, HL, LH, HH e LL, são geradas usando dois filtros, um passa-baixas e outro passa-altas. Como estamos interessados em obter o gradiente das imagens em diferentes resoluções utilizamos uma técnica em que os sinais bidimensionais são decompostos em três bandas que correspondem às filtragens passa-baixas e passa-altas nas direções vertical e horizontal (Mallat e Zhong, 1992a; Mallat e Hwang, 1992; Sun et al., 1994). Nesta implementação são usados três filtros G, L e H, derivados da função spline cúbica, cujos coeficientes são apresentados na Tabela 4.1. Em cada

escala  $2^j$ , o algoritmo decompoõe a imagem  $A_{2^j}f$  em  $A_{2^{j+1}}f$ ,  $D^1_{2^{j+1}}f$  e  $D^2_{2^{j+1}}f$ , que representam respectivamente, as imagens de baixa resolução (banda LL), gradiente na direção vertical (banda LH) e gradiente na direção horizontal (banda LH) na escala  $2^{j+1}$ . Os filtros G e L são usados para calcular as bandas LH e HL e o filtro H é usado para calcular a banda LL. Note que, por conveniência, continuamos a usar a mesma nomenclatura para as sub-imagens LL, LH e HL, embora na nossa implementação a banda HH não seja gerada.

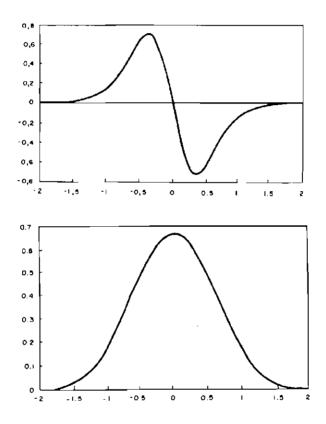


Fig. 4.3 - Função wavelet spline cúbica  $\psi$  e a correspondente função de suavização  $\phi$ .

Para resolver o problema da filtragem nas bordas da imagem, consideramos as últimas linhas e colunas como uma continuação das primeiras linhas e colunas, respectivamente.

Na decomposição waveletem pirâmide, as bandas  $A_{2^{j+1}}f,\,D^1_{2^{j+1}}f$ e $D^2_{2^{j+1}}f$ são dis-

TABELA 4.1 -FILTROS CORRESPONDENTES A WAVELET SPLINE CÚBICA

n	$g_n$	$l_n$	$h_n$
0	-0.00008	0.00003	0.0625
1	-0.01643	0.00727	0.25
2	-0.10872	0.03118	0.375
3	-0.59261	0.06623	0.25
4	0.59261	0.79113	0.0625
5	0.10872	0.06623	
6	0.01643	0.03118	
7	0.00008	0.00727	
8	<u> </u>	0.00003	

poníveis até o nível de resolução  $2^1$ . Desta forma não temos no nível de resolução  $2^0$  (imagem original) a representação para estas bandas. Para que possamos extrair as feições pontuais (bordas) no nível de resolução  $2^0$ , as bandas  $D_{2^{j+1}}^1f$  e  $D_{2^{j+1}}^2f$  são obtidas da seguinte forma. Sejam G e H os filtros passa-altas e passa-baixas usados na decomposição wavelet. Na escala  $2^0$ , o sinal  $A_1^df$  é decomposto em duas sub-imagens:

$$D_1^1 f = A_1^d f * (H, G)$$

$$D_1^2 f = A_1^d f * (G, H),$$
(4.12)

onde A\*(H,G) denota a convolução separável nas linhas e nas colunas da imagem  $A_1^d f$  com os filtros H e G, respectivamente. Este algoritmo é similar ao usado na a obtenção das bandas  $D_{2^{j+1}}^1 f$  e  $D_{2^{j+1}}^2 f$  em cada nível de resolução, a menos do processo de dizimação que, neste caso, não é realizado. O processo de detecção dos pontos de bordas no nível  $2^0$  segue o mesmo procedimento descrito anterior.

Para calcular os pontos de máximos locais, temos que encontrar os dois pontos vizinhos ao ponto P(x,y),  $P_1$  e  $P_2$ , na direção do vetor gradiente, como mostra a Figura 4.4. Como o ângulo do gradiente pode variar continuamente de 0 a  $2\pi$  radianos, é necessário que se faça uma aproximação porque os dados a serem processados são discretos. Na implementação, consideramos apenas quatro direções indicadas pelas regiões norte-sul (NS), oeste-leste (OL), noroeste-sudeste (SE) e sudoeste-nordeste (NE) que definem as direções médias do gradiente, como mostra a Figura 4.5. Seja a direção do gradiente no ponto  $P(x_o, y_o)$ ,  $\alpha_{2^j}(x_o, y_o)$ , e o valor da tangente deste ângulo,  $tan\alpha_{2^j}(x_o, y_o)$ . Então, os pontos vizinhos de  $P(x_o, y_o)$ ,  $Vp(x_o, y_o)$  são definidos da seguinte forma:

Se 
$$tan_{2^{j}}(x_{o}, y_{o}) \leq -(\sqrt{2} - 1)$$
 ou  $tan_{2^{j}}(x_{o}, y_{o}) > (\sqrt{2} + 1)$  (região NS)  
 $Vp(x_{o}, y_{o}) = \{P(x_{o} - 1, y_{o}), P(x_{o} + 1, y_{o})\}$ 

Se 
$$(\sqrt{2} - 1) < tan_{2^{j}}(x_o, y_o) \le (1 + \sqrt{2})$$
 (região NE)  
 $Vp(x_o, y_o) = \{(x_o + 1, y_o + 1), (x_o - 1, y_o + 1)\}$ 

Se 
$$(1 - \sqrt{2}) < tan_{2^{j}}(x_o, y_o) \le (\sqrt{2} - 1)$$
 (região OL)  
 $Vp(x_o, y_o) = \{(x_o, y_o - 1), (x_o, y_o + 1)\}$ 

Se 
$$(-\sqrt{2}-1) < tan_{2^{j}}(x_{o}, y_{o}) \le (1 \sqrt{2})$$
 (região SE)  
 $Vp(x_{o}, y_{o}) = \{(x_{o}-1, y_{o}-1), (x_{o}+1, y_{o}+1)\}.$ 

Todos os algoritmos, desenvolvidos neste trabalho, foram implementados na linguagem C++, inclusive a transformada wavelet.

#### 4.4 Resultados Experimentais

O algoritmo de detecção de feições pontuais descrito na seção anterior é aplicado para três imagens do sensor TM de regiões diferentes, de tamanho 512x512. Estas imagens correspondem a uma área da cidade de Brasília, de uma região de agricul-

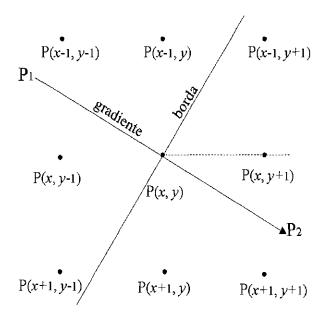


Fig. 4.4 - Vizinhos mais próximos do Ponto  ${\cal P}$  na direção do vetor gradiente.

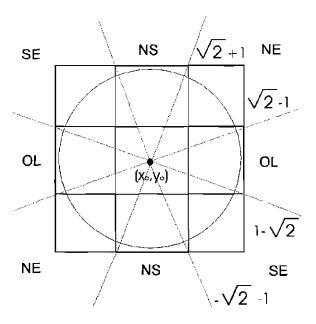


Fig. 4.5 - Direções aproximadas indicadas pelo ângulo  $\alpha.$ 

tura próxima a cidade de Agudos em São Paulo e da região Amazônica mostradas nas Figuras 4.6, 4.7 e 4.8, respectivamente.

As imagens originais são processadas em dois (2) níveis de resolução da transformada wavelet. As Figuras 4.9 , 4.10 e 4.11 mostram as feições pontuais sobrepostas às bandas LL da decomposição wavelet para dois níveis de resolução e as feições pontuais sobrepostas às imagens originais para o nível  $2^0$ , para as imagens originais mostradas nas Figuras 4.6, 4.7 e 4.8, respectivamente. Os valores dos parâmetros usados nos experimentos foram os seguintes:  $\beta=3$  (para eliminar as feições menos significativas),  $T_w=0,95$  para o limiar de contraste e uma vizinhança  $V_p$  de 7x7 pixels na seleção dos máximos locais. Estes mesmos valores foram testados para outras imagens, com sucesso, o que é um indicativo da robustez da metodologia usada.



Fig. 4.6 - Imagem TM da cidade de Brasília, banda  $4.\,$ 

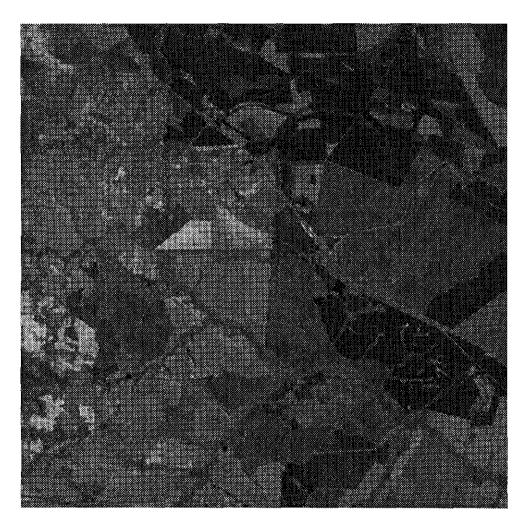


Fig. 4.7 - Imagem TM de uma região de agricultura, banda  $5.\,$ 

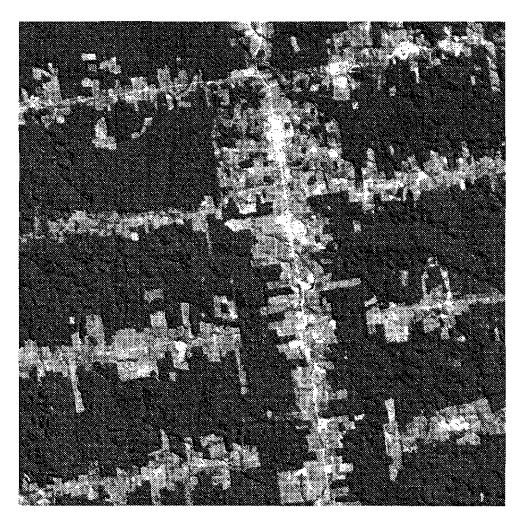


Fig. 4.8 - Imagem TM da região Amazônica, banda  $5.\,$ 

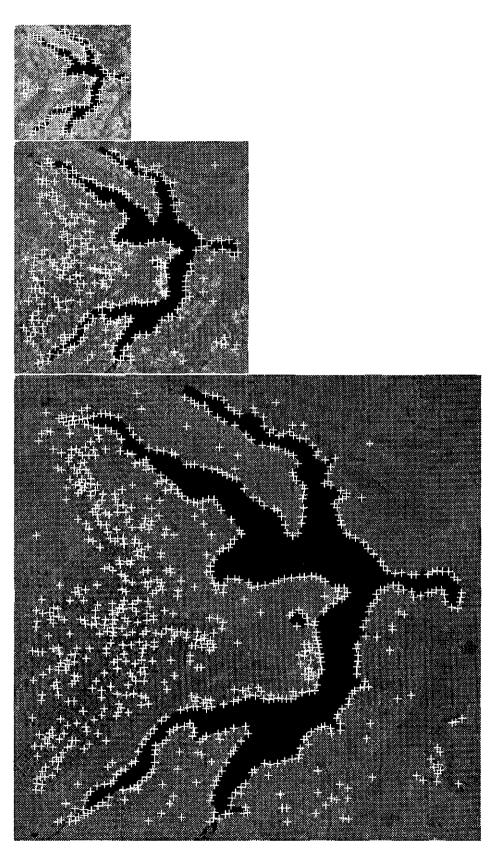


Fig. 4.9 - Feições pontuais: (a) resolução  $2^2$ ; (b) resolução  $2^1$ ; (c) resolução  $2^0$ .

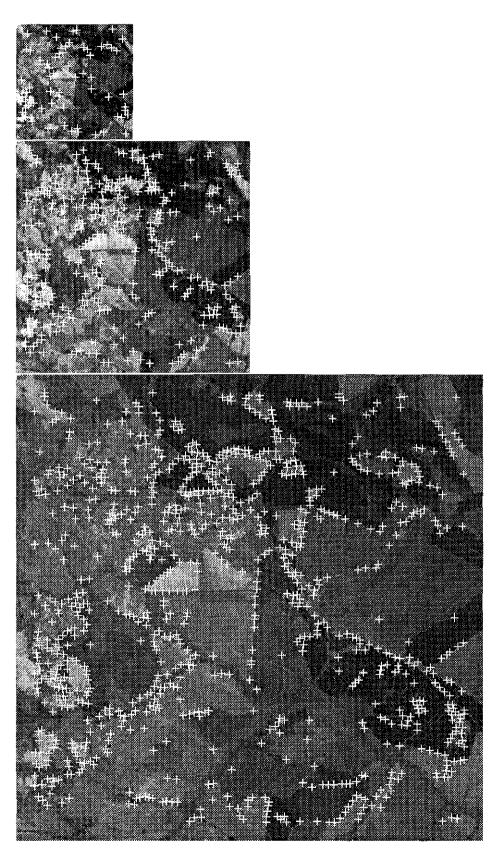


Fig. 4.10 - Feições pontuais: (a) resolução  $2^2;$  (b) resolução  $2^1;$  (c) resolução  $2^0.$ 

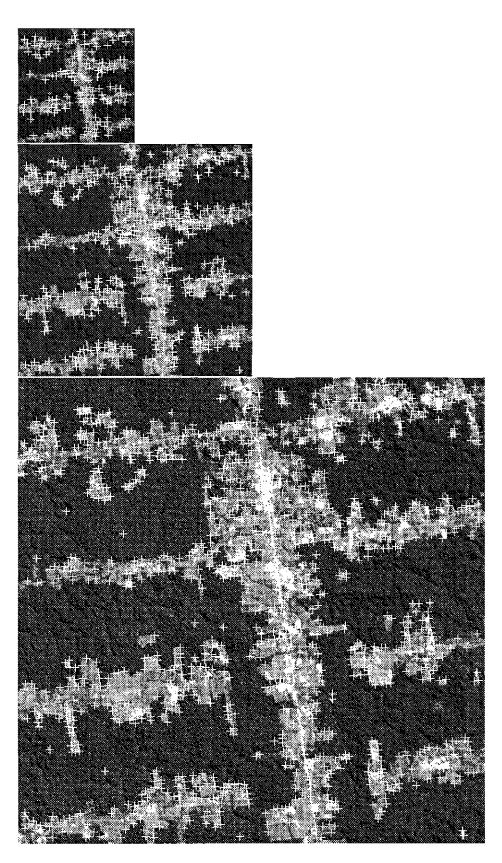


Fig. 4.11 - Feições pontuais:(a) resolução  $2^2$ ; (b) resolução  $2^1$ ; (c) resolução  $2^0$ .

# CAPÍTULO 5

# MEDIDAS DE SIMILARIDADE: ANÁLISE DE DESEMPENHO

No registro de imagens o processo de casamento é uma das fases mais importantes. Métodos de casamento baseados em área usam medidas de similaridade para comparar duas janelas de pontos nas imagens, as quais contém as feições de interesse. As medidas de similaridade indicam o grau de similaridade entre dois padrões a serem casados. A correlação (Pratt, 1974) e a detecção de similaridade seqüencial (Barnea e Silverman, 1972) são as medidas de similaridade mais conhecidas. Porém, outras medidas de similaridade tais como os operadores morfológicos de casamento (Khosravi e Shafer, 1996) e medidas baseadas na norma  $L_1$  (Brunelli e Messelodi, 1995) têm sido propostas.

O desempenho de uma medida de similaridade, no processo de casamento, depende da aplicação, tipo de padrões a serem comparados, nível de ruído nas imagens, etc. Uma forma de avaliar uma medida de similaridade é medindo a sua capacidade de detectar casamentos corretos. Portanto, a obtenção de um número alto de casamentos corretos é desejável, e é um indicativo do bom desempenho de uma medida de similaridade.

Na Seção 5.1 apresentamos uma breve descrição de quatro medidas de similaridade: coeficiente de correlação, uma medida baseada na norma  $L_1$  (L1G), detecção de similaridade seqüencial (DSS) e o operador de casamento morfológico (GHMT). A escolha destes métodos foi ditada por duas razões:1) a correlação e o DSS são métodos tradicionais e muito usados; 2) os métodos GHMT e L1G têm mostrado, em algumas situações, bom desempenho quando comparados ao método de correlação (Brunelli e Messelodi, 1995; Khosravi e Shafer, 1996).

No processo de casamento, a situação ideal é aquela em que o método obtém o número máximo de casamentos corretos e nenhum casamento errado. Vamos ver na Seção 5.2 que medidas de confiança podem ser usadas, junto com as medidas de

similaridade, para reduzir o número de casamentos errados. A Seção 5.3 apresenta uma avaliação do desempenho das medidas de similaridade acima mencionadas.

#### 5.1 Medidas de Similaridade

Antes de prosseguir com as definições dos métodos de similaridade, vamos introduzir algumas notações e definições usadas nesta seção.

Sejam duas imagens f e g em níveis de cinza. Seja t uma janela de pontos, extraída da imagem g, com L linhas e K colunas, centrada no ponto (m,n), a qual chamaremos de padrão. Seja  $W = \{(i,j): i=0,...,L-1; j=0,...,K-1\}$  o suporte de t. A comparação entre o padrão t e uma janela de f, com suporte W e centrada em (x,y), é realizada através de medidas de similaridade.

Para simplificar, a referência à uma determinada medida é feita através de abreviaturas (entre parênteses) da seguinte forma: coeficiente de correlação (CORREL), detecção de similaridade sequencial (DSS), operador de casamento morfológico Gray Hit-or- $Miss\ Transform\ (GHMT)\ e\ medida\ baseada\ na\ norma\ L_1\ (L1G).$ 

#### 5.1.1 Coeficiente de Correlação

Neste método, o casamento é realizado através da maximização do coeficiente de correlação entre o padrão t e a imagem f definido por

$$C(x,y) = \frac{1}{w_c^2} \frac{\sum_{(i,j)\in W} \{f(x+i,y+j) - \mu_f\} \{t(i,j) - \mu_t\}}{[\sum_{(i,j)\in W} \{f(x+i,y+j) - \mu_f\}^2 \{t(i,j) - \mu_t\}^2]^{1/2}},$$
(5.1)

onde  $\mu_f$  e  $\mu_t$  são as médias locais de f e t, respectivamente, W é o suporte de t e  $w_c^2$  é a cardinalidade do suporte.

O coeficiente de correlação é uma das medidas de similaridade mais tradicionais. Alguns autores consideram o coeficiente de correlação uma técnica de casamento confiável e por isso tem sido muito usada (Venkateswarlu e Chatterjee, 1992; Wacker et al., 1983).

#### 5.1.2 Medida de Similaridade baseada na Norma $L_1$

Brunelli e Messelodi (1995) avaliaram o desempenho de algumas medidas de similaridade baseadas nas normas  $L_1$  e  $L_2$  para reconhecimento de faces e capas de livros. Nos experimentos, as medidas baseadas na norma  $L_1$  mostraram desempenho superior às medidas baseadas na norma  $L_2$ .

Uma destas medidas de similaridade usada por Brunelli e Messelodi (1995) é

$$G(x,y) = 1 - \frac{\sum_{(i,j)\in W} |f'(x+i,y+j) - t'(i,j)|}{\sum_{(i,j)\in W} |f'(x+i,y+j)| + |t'(i,j)|}$$
(5.2)

onde t' e f' são os valores de t e f normalizados para média 0 e variância 1.

Os valores de G estão normalizados no intervalo [0,1], de tal forma que

$$G(x,y) = 1 \Leftrightarrow t'(i,j) = f'(x+i,y+j),$$
  
$$G(x,y) = 0 \Leftrightarrow t'(i,j) = -f'(x+i,y+j),$$

para 
$$(i, j) \in W$$
.

Similar ao coeficiente de correlação, valores altos de G é indicativo de maior grau de similaridade entre os padrões.

### 5.1.3 Detecção de Similaridade Sequencial

O método de detecção de similaridade seqüencial (DSS) proposto por Barnea e Silverman (1972) é computacionalmente muito mais simples que os métodos de correlação. Ele é baseado na soma das diferenças absolutas entre os valores de f e t,

$$S'(x,y) = \sum_{(i,j)\in W} |f(x+i,y+j) - t(i,j)|, \qquad (5.3)$$

e no seguinte método de busca sequencial. Os valores das diferenças absolutas, dentro do suporte W, são acumulados até que um valor de limiar pré-determinado seja excedido; o número de pontos examinado quando este limiar é excedido é armaze-

nado como um avaliador do teste (taxa); quando todas as sub-imagens de f tiverem sido examinadas, a sub-imagem com a maior taxa e menor erro é considerada o melhor casamento.

Para diminuir a sensibilidade às diferenças espectrais entre as duas imagens, podese calcular a soma das diferenças absolutas dos valores dos pixels relativos as suas médias dentro das janelas

$$S(x,y) = \sum_{(i,j)\in W} |\{f(x+i,y+j) - \mu_f\} - \{t(i,j) - \mu_t\}|_{\perp}.$$
 (5.4)

Neste trabalho, a busca sequencial é realizada da seguinte mancira. Dados um padrão t e uma sub-imagem  $f_i$  na imagem de ajuste calcula-se a medida de similaridade, conforme a Equação 5.3, considerando todos os pontos dentro da janela de suporte W. O valor desta medida é usado como o valor de limiar inicial,  $T_0$ . Para a próxima janela  $f_j$ ,  $j \neq i$ , os valores das diferenças absolutas (Equação 5.3), dentro do suporte W, são acumulados ( $T_j$ ) até que  $T_j > T_0$ ; o número de pontos examinados quando  $T_0$  é excedido é armazenado como  $N_p$ ; quando todas as sub-imagens de f tiverem sido examinadas, a sub-imagem com maior  $N_p$  e menor erro,  $T_j$ , é considerada o melhor casamento. Note-se que os valores de  $N_p$  e  $T_j$  são atualizados a medida que um casamento melhor do que o anterior é encontrado.

#### 5.1.4 Operador de Casamento Morfológico: GHMT

A morfologia matemática foi iniciada por George Matheron e Serra, que construíram duas classes de operadores elementares: as dilatações e as erosões (Serra, 1982). A idéia de Matheron (1967) era decompor operadores crescentes invariantes à translação, entre imagens binárias, em termos de dilatações e crosões. Banon e Barrera (1991) estenderam esta idéia para quaisquer operadores invariantes em translação (não necessariamente crescentes), usando mais duas classes de operadores elementares: anti-dilatações e anti-erosões. Estes operadores são chamados operadores elementares porque a partir da composição deles é possível representar outros operadores (Banon e Barrera, 1991, 1993). Uma abordagem completa sobre a morfologia

matemática pode ser encontrada em Serra (1982); Serra (1988); Heijmans (1994); Banon (1995); Banon e Barrera (1998a) e Banon (1998b).

Khosravi e Shafer (1996) introduziram um operador morfológico que pode ser usado para localizar um padrão dentro de um sinal em níveis de cinza. Este operador, denominado GHMT (*Gray Hit-or-Miss Transform*), usa a soma de duas erosões em níveis de cinza.

Seja  $\psi_{f,t}^e$  a operação de erosão em níveis de cinza da imagem f por um padrão t. A comparação entre f e t é realizada através da maximização do operador morfológico GHMT definido por

$$M(x,y) = \psi_{f,t}^{e}(x,y) + \psi_{-f,-t}^{e}(x,y). \tag{5.5}$$

A erosão morfológica em níveis de cinza de uma imagem f pelo elemento estruturante (padrão) t é definido como

$$\psi_{f,t}^{e}(x,y) = \min_{(i,j) \in W} \{ f(x+i,y+j) - t(i,j) \}, \tag{5.6}$$

Substituindo a Equação 5.6 na Equação 5.5 temos que

$$M(x,y) = \min_{(i,j)\in W} \{ f(x+i,y+j) - t(i,j) \} + \min_{(i,j)\in W} \{ -f(x+i,y+j) + t(i,j) \}.$$
(5.7)

O operador GHMT pode ser reescrito da seguinte forma

$$M(x,y) = \min_{(i,j)\in W} \{f(x+i,y+j) - t(i,j)\} - \max_{(i,j)\in W} \{f(x+i,y+j) - t(i,j)\}.$$
(5.8)

Da expressão (5.8) tem-se que os valores de GHMT são não-positivos e que, o valor máximo é zero e ocorre somente quando f(x+i,y+j) e t(i,j) são idênticos para  $(i,j) \in W$ . O operador GHMT é insensível às variações nas médias das imagens

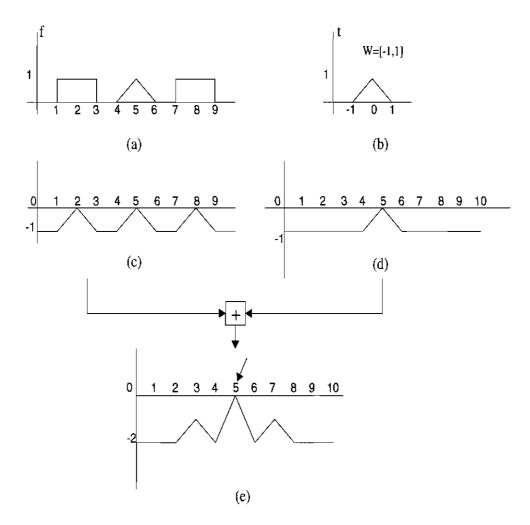


Fig. 5.1 - Aplicação do operador GHMT sobre o sinal unidimensional f: (a) sinal f; (b) padrão t; (c) resultado da operação de erosão de f por t; (d) resultado da operação de erosão de -f por -t; (e) o símbolo  $\checkmark$  indica a posição onde se deu o melhor casamento.

FONTE: modificada de Khosravi e Schafer (1996).

e mais rápido (aproximadamente três vezes) que o critério de correlação pois usa comparações ao invés de multiplicações.

A Figura 5.1 ilustra as operações de erosões de um sinal unidimensional f pelos padrões  $t \in -t$ , e o resultado da aplicação do operador GHMT sobre este sinal.

### 5.2 Medidas de Confiança

Quando, por exemplo, usamos o coeficiente de correlação para comparar duas feições, a maximização da medida para a escolha do melhor casamento não garante que o casamento obtido seja correto. Embora o valor do coeficiente seja máximo, valores pequenos do coeficiente de correlação pode implicar em um casamento errado. Dizemos, neste caso, que este casamento tem baixa credibilidade ou confiança. Problemas tais como ruído, oclusão e baixo contraste nas imagens podem levar ao casamento incorreto. Então, é necessário que haja alguma garantia de que os pontos casados sejam corretos.

Para resolver as ambigüidades no processo de casamento, medidas de confiança podem ser usadas. Medidas de confiança são funções que atribuem uma "nota" ao par de pontos casados, indicando se aquele par de pontos casados é confiável ou não. Se esta nota estiver dentro de uma faixa de aceitação dizemos que o casamento tem alta credibilidade e então é declarado correto. Do contrário, ele é declarado errado e é descartado.

Uma medida de confiança usada para a técnica DSS e definida em Venkaterswalu e Chatterjee (1992) é:

$$MC(x_o, y_o)' = \frac{S(x_o, y_o)}{[1/(K^2 - 1)] \sum_{i,j=0}^{K-1} S(x_o + i, y_o + j)}.$$
 (5.9)

 $S(x_o, y_o)$  é o valor mínimo da medida de similaridade, K é o tamanho de uma janela quadrada centrada em  $(x_o, y_o)$ . O somatório é realizado na vizinhança do ponto para o qual o erro é mínimo. Se o ponto  $(x_o, y_o)$  indica a posição em que o padrão se casa melhor com a imagem, é desejável que, nesta posição, o valor de MC' seja pequeno

(< 1). Valores altos de S na vizinhança do ponto  $(x_o, y_o)$  e em média maiores do que o valor  $S(x_o, y_o)$ , garantem uma maior confiabilidade no ponto de casamento.

Como os valores da Equação 5.9 diminuem à medida que a confiabilidade do ponto de casamento aumenta, a nomenclatura "medida de confiança" pode provocar uma certa confusão na sua interpretação. A interpretação de que altos valores de uma medida de confiança indicam maior confiança do casamento parece ser mais coerente. Para que esta coerência seja garantida, propomos como alternativa a seguinte medida de confiança

$$MC(x_o, y_o) = 1 - \frac{S(x_o, y_o)}{[1/(K^2 - 1)] \sum_{i,j=0}^{K-1} S(x_o + i, y_o + j)}.$$
 (5.10)

Desta forma, valores positivos e altos de MC indicam maior confiança no ponto de casamento e vice-versa. Similarmente, podemos usar esta mesma medida para o método GHMT, considerando o módulo de GHMT.

Para os métodos de correlação e L1G, o próprio valor da medida, que é normalizado, pode ser usado para medir a confiabilidade de um casamento. Neste caso, valores altos de correlação e L1G indicam um maior grau de similaridade entre os pontos e, portanto, maior o grau de confiança do casamento.

Quando, no processo de casamento, usamos uma medida de confiança duas situações podem ocorrer:

- 1) casamentos errados, denominados alarmes falsos, são eliminados;
- 2) casamentos corretos são eliminados, os quais são denominados casamentos omissos;

Uma boa medida de confiança é aquela que obtém um número mínimo de casamentos omissos e alarmes falsos no processo de casamento, mantendo assim uma alta porcentagem de casamentos corretos. As medidas de similaridade descritas nas Seção 5.1 são implementadas, e seus desempenhos são analisados na próxima seção.

## 5.3 Avaliação dos Critérios de Similaridade

Para a avaliação de medidas de similaridade várias metodologias têm sido usadas (Venkaterswalu e Chatterjee, 1992; Brunelli e Messelodi, 1995; Khosravi e Shafer, 1996). Venkaterswalu e Chatterjee (1992) analisam o desempenho dos métodos DSS e correlação para imagens do espectro infra-vermelho e de baixo contraste. Brunelli e Messelodi (1995) investigam o desempenho de métodos de similaridade no reconhecimento de faces e capas de livros. Khosravi e Shafer (1996) testam o desempenho dos métodos de correlação e GHMT na presença de ruídos gaussianos e impulsivos.

Neste trabalho, vamos investigar o desempenho dos métodos CORREL, L1G, DSS e GHMT no processo de casamento para dados reais de imagens de satélites. Para a avaliação, selecionamos seis pares de imagens multi-temporais, de tamanho 512x512 pixels: duas imagens de áreas urbanas, duas de florestas e duas imagens de agricultura. A escolha deste conjunto de imagens testes foi devida à diversidade de informações contidas nelas (textura, diferenças de contraste, etc) que, no nosso parecer, permite uma avaliação robusta. Estes dados testes se diferem dos dados usados por Brunelli e Messelodi (1995) por possuirem muitos detalhes e serem de diferentes datas, conferindo assim um maior grau de dificuldade no processo de casamento. Todas as imagens, com exceção dos pares de imagens da região de agricultura, foram realçadas antes do processo de casamento. A técnica de realce usada foi o realce de histograma linear.

Em todos os experimentos descritos na próxima seção, as imagens de referência e de ajuste são decompostas (decomposição wavelet) em dois níveis de resolução e o processo de casamento é realizado no nível de resolução mais baixo. Sejam  $f_r$  e  $f_s$  as imagens de referência e de ajuste, no nível de resolução mais baixo. Dois conjuntos de feições são extraídos nas imagens  $f_r$  e  $f_s$  através do algoritmo apresentado no Capítulo 4. As imagens  $f_r$  e  $f_s$ , juntamente com estas feições, são usadas para realizar o processo de casamento. Os experimentos são realizados para todos os seis pares de imagens, e os resultados finais são dados pela média dos valores obtidos

para cada par de imagens teste.

Nas próximas seções apresentamos três experimentos para avaliação de desempenho dos métodos descritos na Seção 5.1, com e sem medida de confiança. Os procedimentos para realização dos experimentos foram definidos de forma a simular as condições reais em que o algoritmo de registro proposto nesta tese é usado.

#### 5.3.1 Experimento 1: desempenho em função do tamanho da janela

Neste experimento, o objetivo é comparar o desempenho dos métodos CORREL, L1G, DSS e GHMT, considerando diferentes tamanhos de janelas e sem medida de confiança.

O processo de casamento, sem medida de confiança, é realizado da seguinte forma. Para cada par de imagens, dez  $(N_t=10)$  pares de casamentos corretos são selecionados. Para cada janela quadrada de MxM pontos centrada em uma das  $N_t$  feições na imagem de referência (padrão), procura-se a janela de mesmo suporte centrada em uma das feições na imagem de ajuste, que seja a mais similar ao padrão de acordo com o critério de similaridade escolhido.  $N_t$  pares de pontos casados são encontrados. Dentre os  $N_t$  pares de pontos casados, alguns são casamentos corretos  $N_c$  e os restantes são casamentos errados  $(N_{te}=N_t-N_{tc})$ , os quais chamaremos de alarmes falsos. Alarmes falsos são casamentos errados mas que foram declarados corretos no processo de casamento. Quanto maior a porcentagem de casamentos corretos melhor é o desempenho da medida de similaridade. A taxa de acertos (TA), neste caso, é definida pela razão entre o número de pontos casados corretamente e o número de pontos pré-definidos como corretos, ou seja  $N_{tc}/N_t$ .

A Figura 5.2 mostra o gráfico da taxa média de acertos (em porcentagem) de cada método, em função do tamanho da janela de casamento. Podemos observar neste experimento que, para os métodos CORREL, L1G e DSS, o desempenho aumenta à medida que o tamanho da janela aumenta. De forma diferente, o método GHMT apresenta algumas oscilações, com um pico saliente em torno de M=13. Este comportamento pode indicar uma certa sensibilidade do método GHMT com relação à

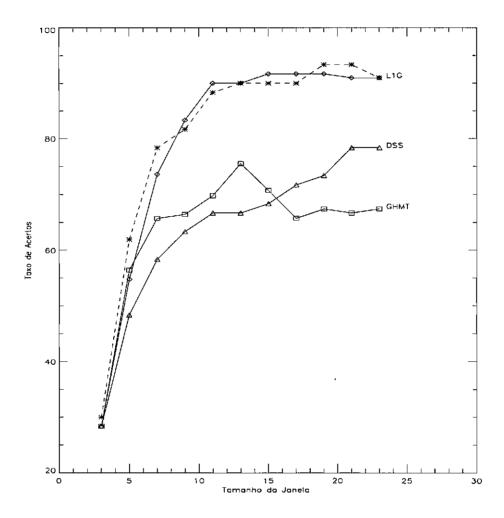


Fig. 5.2 - O desempenho médio do método CORREL (linha tracejada) é comparado com o desempenho médio dos métodos L1G, DSS e GHMT, considerando diferentes tamanhos da janela de casamento.

variação do tamanho da janela. Os métodos CORREL e L1G apresentam desempenhos semelhantes e superiores aos outros métodos, principalmente para janelas maiores do que 9.

## 5.3.2 Experimento 2: desempenho com medida de confiança

Neste experimento, fazemos uma avaliação dos quatro métodos, considerando medidas de confiança. O processo de casamento é realizado para todas as N feições na imagem de referência. Para cada janela quadrada de MxM pontos centrada em uma das N feições em  $f_r$  (padrão), procura-se a janela de mesmo suporte centrada em uma das feições em  $f_s$ , que seja a mais similar ao padrão de acordo com o critério de similaridade escolhido. Dos N casamentos, obtidos sem o uso de medida de confiança,  $N_c$  são corretos e  $N_e$  são casamentos errados. Atribui-se a cada casamento um valor de medida de similaridade que é utilizado para inferir se o casamento é correto ou falso.

Com a inclusão da medida de confiança no processo de casamento, os cálculos das taxas de acertos, alarmes falsos e omissos são realizadas da seguinte maneira. Dados os N casamentos obtidos sem medida de confiança, somente os casamentos cujas medidas de confiança estão dentro de uma faixa de aceitação são armazenados. Os casamentos restantes são considerados de baixa confiança e são, então, eliminados. Assim, dos  $N_c$  casamentos corretos alguns são eliminados por apresentarem baixa confiança apesar de serem casamentos corretos. Estes casamentos são classificados como casamentos omissos  $(N_{ce})$ . A outra parte é armazenada e declarada como casamentos corretos  $(N_{cc})$ . Da mesma forma, dos  $N_c$  casamentos errados alguns são corretamente eliminados  $(N_{ee})$  e outros são incorretamente declarados como corretos  $(N_{ec})$ , que são os chamados alarmes falsos.

O uso de uma medida de confiança cada vez mais restritiva faz com que o número de alarmes falsos diminua e que a taxa de casamentos omissos aumente. A razão entre o número de casamentos omissos  $(N_{ce})$  e o número total de casamentos obtidos (N) é chamada de taxa de omissos  $(N_{ce}/N)$ . A razão entre o número de alarmes falsos

 $(N_{ec})$  e N é chamada de taxa de alarmes falsos  $(N_{ec}/N)$ .

Os valores de S na Equação 5.4, e o módulo de M na Equação 5.8 são comparáveis no sentido que, a busca do melhor casamento é realizada pela minimização destas medidas. Por outro lado, os métodos CORREL e L1G exibem um comportamento oposto, onde a busca do melhor casamento é obtido pela maximização das respectivas medidas.

O nosso objetivo é comparar os quatro métodos considerando a mesma medida de confiança, definida na Equação 5.10. No cálculo da medida de confiança, consideraremos as seguintes medidas de erro:

$$E^{2}(x,y) = 1 - C(x,y). (5.11)$$

$$E^{1}(x,y) = 1 - G(x,y), (5.12)$$

para os métodos CORREL e L1G, respectivamente, onde C é o coeficiente de correlação definido na Equação 5.1 e G é a medida norma  $L_1$  definida na Equação 5.2. Com estas modificações, os valores das medidas |M|,  $E^1$  e  $E^2$  são usados para calcular a função MC (Equação 5.10) para cada um destes métodos de forma similar ao método DSS.

Neste experimento, o processo de casamento é realizado para os seis pares de imagens, para janelas de tamanho M=13. Para cada ponto de feição na imagem de referência procura-se o ponto de feição mais similar na imagem de busca, usando cada uma das quatro medidas de similaridade. No ponto de melhor casamento, calcula-se a medida de confiança correspondente ao método. Se o valor da medida de confiança, para este casamento, estiver dentro da faixa de aceitação (confiança), o casamento é aceito como correto, do contrário ele é eliminado. Para a análise dos resultados, o número de casamentos corretos, a taxa de alarmes falsos e a taxa de omissos são calculados.

Os gráficos nas Figuras 5.3(a)-(c) apresentam, para os quatro métodos, o número

médio de casamentos corretos (# casamentos corretos), as taxas (%) médias de alarmes falsos e de omissos em função dos valores das medidas de confiança, respectivamente. Pode-se observar que, para um mesmo valor de medida de confiança (MC), o método CORREL é o que apresenta a maior taxa de alarmes falses e menor taxa de omissos. A baixa taxa de omissos faz com que o método CORREL detecte um maior número de casamentos corretos, seguido do método L1G. O desempenho de cada um dos métodos pode ser alterado de acordo com a escolha do valor MC.

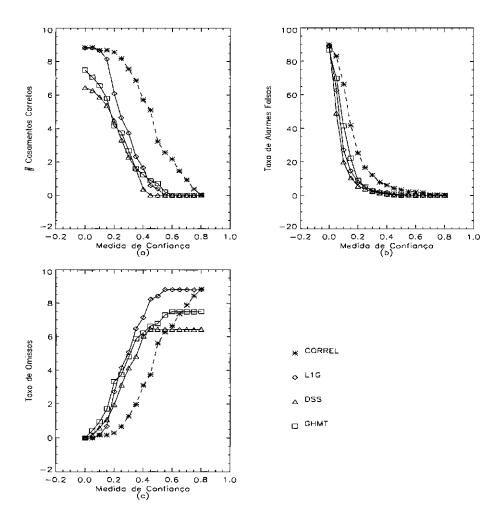


Fig. 5.3 - Desempenho médio dos métodos CORREL, L1G, DSS e GHMT com medidas de confiança e tamanho fixo da janela de casamento (M=13).

Vamos considerar que um bom desempenho seja aquele em que o método consiga um número de casamentos corretos alto com uma baixa taxa de falsos alarmes. O gráfico na Figura 5.4 mostra como o número de casamentos corretos varia em função da

taxa de alarmes falsos, para os quatro métodos. Pode-se observar que os métodos de correlação e L1G apresentam melhores desempenhos que os métodos DSS e GHMT para valores de  $N_{cc}$  maiores do que aproximadamente 3. Na Figura 5.3(a) podemos observar que, os valores de MC (MC < 0.3, aproximadamente) para os quais os métodos DSS e GHMT possuem um número maior de casamentos corretos é justamente aqueles para os quais a taxa de alarmes falsos é mais alta. O contrário ocorre para o método CORREL, que para valores de MC maiores (MC > 0.3, aproximadamente) obtém-se um número alto de casamentos corretos com uma baixa taxa de alarmes falsos. Neste experimento, podemos concluir que o método CORREL e L1G apresentam desempenhos similares, com uma leve vantagem do método CORREL quando  $N_{cc} > 4$ .

Podemos observar que a curva na Figura 5.4 é similar às curvas ROC (Receiver Operating Characteristics), usadas por exemplo em sistemas de radar quando se quer decidir se o alvo foi detectado ou não. Estas curvas plotam a probabilidade de detecção  $(P_d)$  versus a probabilidade de alarmes falsos  $(P_f)$ .  $P_d$  indica a probabilidade do radar detectar o alvo quando o alvo está presente e  $P_f$  indica a probabilidade do radar detectar o alvo quando o alvo não está presente (Trees, 1968).

# 5.3.3 Experimento 3: desempenho dos métodos CORREL e L1G com medida de confiança

Neste experimento fazemos uma avaliação do desempenho somente dos métodos CORREL e L1G, considerando como medida de confiança o próprio valor da medida de similaridade no ponto de casamento. O procedimento usado é similar ao do experimento anterior.

Os gráficos nas Figuras 5.5(a)-(c) apresentam o número médio de casamentos corretos (# casamentos corretos), as taxas (%) médias de alarmes falsos e de omissos em função dos valores das medidas de confiança, respectivamente, para os métodos CORREL e L1G. De um modo geral, o método CORREL apresenta maior taxa de alarmes falsos e menor taxa de omissos, para os mesmos valores de medidas de

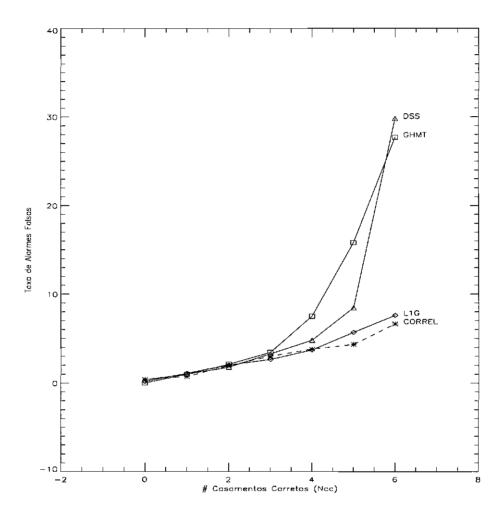


Fig. 5.4 - Taxa de alarmes falsos em função do número de casamentos corretos para os métodos CORREL, L1G, DSS e GHMT (M=13).

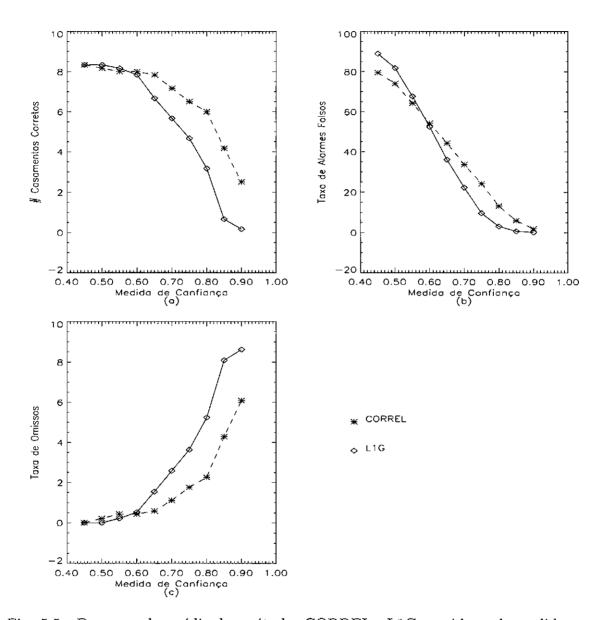


Fig. 5.5 - Desempenho médio dos métodos CORREL e L1G considerando medidas de confiança e tamanho fixo da janela de casamento (M=13).

confiança. Por outro lado, quando fixamos o número de casamentos corretos podemos observar que o método CORREL apresenta uma taxa de alarmes falsos menor do que o método L1G à partir de um certo valor de  $N_{cc}$ , como pode-se observar no gráfico da Figura 5.6. Ou seja, os valores de MC para os quais o número de casamentos corretos para o método L1G é igual ao do método CORREL, correspondem a taxas de alarmes falsos um pouco maiores do que para o método CORREL. Novamente, a superioridade do desempenho do método CORREL sobre o método L1G é verificada.

#### 5.3.4 Conclusão

No algoritmo, que será apresentado no Capítulo 6, os parâmetros da função de transformação são estimados no nível de resolução mais baixa, que por sua vez são refinados nos níveis de resolução mais alta. Neste nível de processamento a precisão dos pontos de controle não é o fator mais importante, pois o que se deseja é ter uma estimativa, ainda que grosseira, do modelo de distorção. O que se deve ser evitado, nesta fase, é a presença de um número grande de casamentos falsos que possam levar à uma estimativa do modelo incorreta, que se propagará nos níveis de processamento subsequentes. Portanto, deve-se eliminar os casamentos falsos antes do cálculo da função de transformação inicial. Este processo de eliminação dos casamentos falsos é tanto mais fácil quanto menor for o seu número.

Além disso, a fase de casamento inicial é a que demanda mais tempo, pois o processo de busca do melhor casamento é mais exaustivo. Comparado ao método de correlação o tempo computacional dos métodos L1G, DSS e GHMT é aproximadamente 1,6,0,6 e 0,3, respectivamente. Como o método de correlação apresentou melhor desempenho em relação aos métodos DSS e GHMT e um pouco melhor do que o método L1G, ele é o que melhor atende os requisitos exigidos pelo algoritmo de registro proposto neste trabalho.

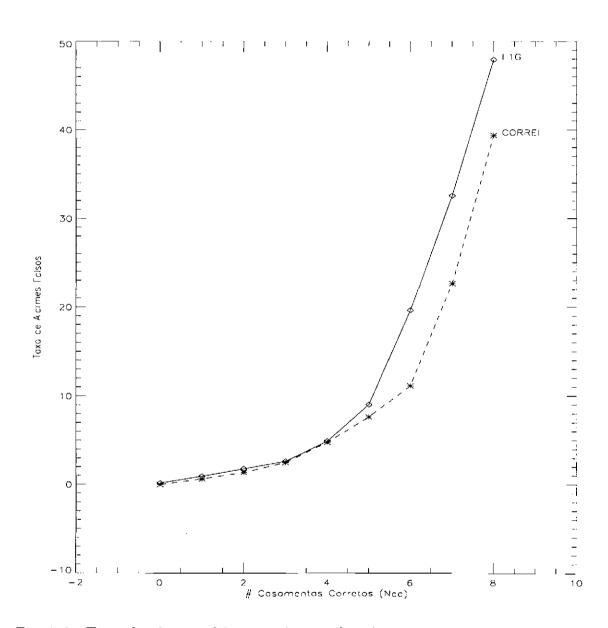


Fig. 5.6 - Taxa de alarmes falsos em função do número de casamentos corretos para os métodos CORREL e L1G (M=13).

# CAPÍTULO 6

# ALGORITMO DE REGISTRO EM MÚLTIPLAS RESOLUÇÕES

Neste Capítulo se apresenta um novo método de registro de imagens de sensoriamento remoto, baseado na análise em multi-resolução através da transformada wavelet.

A representação em multi-resolução fornece um esquema hierárquico que permite interpretar a informação da imagem em diferentes níveis de resolução. Na resolução mais baixa os detalhes da imagem são representados por poucas amostras. Desta forma, é natural primeiramente analisar uma imagem a uma resolução mais baixa, e então continuar a análise nas resoluções mais altas. Este procedimento é chamado de estratégia de processamento *coarse-to-fine*, e possui baixa complexidade computacional.

Como o problema de registro de imagens envolve um esforço computacional relativamente alto, a estratégia coarse-to-fine se mostra bem adequada a este problema e, por isso é utilizada neste trabalho. Para prover a estrutura coarse-to-fine usamos a transformada wavelet que permite:

- análise no domínio do espaço pixel-a-pixel (banda de baixas freqüências);
- análise no domínio da freqüência banda a banda (bandas de altas freqüências);
- realce de feições significativas no nível de resolução baixa com atenuação das feições fracas;
- localização de um número de feições na imagem (pontos de bordas, contornos dos objetos) consistentes nos diferentes níveis de resolução;
- manipulação das imagens em diferentes resoluções;
- processamento da imagem na resolução baixa (redução do ruído e processamento rápido);

- refinamento dos parâmetros da função de transformação (precisão de registro);
- redução do espaço de busca: casamento dos objetos em determinado nível ocorre dentro da área de sobreposição determinada no nível de resolução anterior;
- implementação simples, eficiente, de baixo custo computacional comparada a de métodos tradicionais de busca pixel-a-pixel.

O método de registro proposto se baseia na hipótese de que as imagens possuem respostas espectrais semelhantes, e as diferenças de rotação e escala entre elas são pequenas. O problema de escala é resolvido através da reamostragem das imagens para um mesmo tamanho de pixel. No caso das imagens de satélites o problema da rotação não é relevante, considerando-se que elas são normalmente alinhadas e quando não o são, a informação da diferença de rotação pode ser obtida. Resultados experimentais, apresentados no Capítulo 7, mostram que o algoritmo é capaz de registrar imagens com uma diferença de rotação de até 20 graus.

A Figura 6.1 mostra os estágios envolvidos no algoritmo de registro proposto. O processo de extração de feições nas imagens de ajuste e de referência é realizado automaticamente através da seleção de alguns pontos de borda significativos nas imagens, usando os máximos locais da transformada wavelet. Estes pontos são então usados no processo de casamento pontual. Em cada nível de resolução da decomposição wavelet três sub-imagens, uma de baixa freqüência (LL) e duas de alta freqüência (LH e HL), são usadas nos processos de extração e casamento de feições pontuais. O resultado do registro obtido na resolução mais baixa  $2^j$  é, então, usado na próxima resolução mais alta  $2^{j-1}$  criando, assim, uma estrutura dentro da qual todas as fases do processamento se interagem.

O processo de casamento de feições é realizado usando o método de correlação e inclui uma etapa de verificação de consistência dos casamentos para eliminar os casamentos falsos, inevitáveis na fase inicial. Este procedimento garante uma estimativa correta

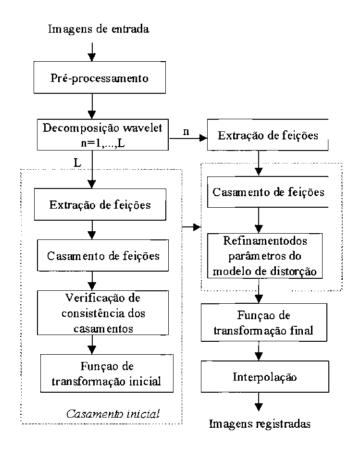


Fig. 6.1 - Estágios de processamento envolvidos no algoritmo de registro.

da função de transformação. Na fase seguinte do processamento, os parâmetros da função de transformação são refinados nos níveis de resolução progressivamente mais fina, seguindo a estratégia *coarse-to-fine*. Após todos os níveis da pirâmide terem sido processados, os parâmetros obtidos são usados para transformar a imagem de ajuste. O interpolador bilinear é usado para estimar os níveis de cinza da imagem de ajuste nas posições determinadas pela função de transformação.

No Capítulo 7 são apresentados vários resultados experimentais que verificam a eficiência computacional e o desempenho do algoritmo proposto. Nas seções seguintes, descrevemos com detalhes cada uma das etapas envolvidas no processo de registro.

#### 6.1 Pré-processamento

Na fase de pré-processamento as imagens a serem registradas são preparadas para serem registradas.

Em relação à escala das imagens são considerados dois casos no processo de registro:

- as imagens são de mesmo sensor e possuem o mesmo tamanho de pixel;
- as imagens são de diferentes sensores e possuem diferentes tamanhos de pixel.

No caso de as imagens possuirem o mesmo tamanho de pixel, nenhum procedimento de uniformização de escala é realizado. Caso contrário, antes do registro, as imagens são reduzidas a uma mesma escala através da reamostragem. Técnicas de interpolação ou de restauração podem ser usadas para reamostrar as imagens (Niblack, 1986; Câmara, 1983; Fonseca, 1988; Fonseca et al.,1993)

Antes do processo de registro, as imagens podem ser processadas para se realçarem suas feições e assim facilitar as suas extrações. Vários tipos de filtros, tais como os filtros lineares passa-baixas, de mediana e morfológicos têm sido usados para reduzir o ruído ou textura nas imagens. Em Fonseca et al. (1998) testamos uma classe de filtros chamados filtros Peer Group Average (PGA) (Hewer, 1997). Estes filtros suavizam as imagens e preservam a informação de bordas, que é muito importante

no processo de detecção de feições. Hewer et al. (1997) fazem uma comparação favorável destes filtros em relação ao filtro da mediana.

As imagens de satélites, geralmente, possuem baixo contraste e então é conveniente realçar o contraste destas imagens antes da seleção dos pontos de controle. No caso do registro de imagens de radares a fase de pré-processamento para reduzir o efeito do ruído speckle é muitas vezes necessária. O problema central está na escolha do filtro mais adequado para uma dada aplicação. Alguns pesquisadores têm realizado estudos de avaliação do desempenho de diversos filtros de redução do ruído speckle e tentado caracterizar os problemas inerentes a cada um dos métodos (Santa'Anna, 1995). Neste trabalho, o filtro usado para a redução de ruído speckle é o filtro gama, que usa uma variante do método de otimização iterativo simulated annealing para ajustar um modelo da secção transversal do alvo (modelo localmente constante) para os dados de imagem (Oliver e Quegan, 1998). O resultado deste filtro é uma imagem suavizada, mas com os detalhes da imagem original preservados (ENVI, 1999).

#### 6.2 Detecção de Feições Pontuais

Depois de reduzir as imagens para uma mesma resolução espacial, a decomposição wavelet das imagens de ajuste  $f_s$  e de referência  $f_r$  em L níveis é calculada, resultando nas sub-imagens

$$\left( A_{2^L}^d f_r, \, (D_{2^j}^1 f_r)_{1 \le j \le L}, \, (D_{2^j}^2 f_r)_{1 \le j \le L} \right)$$

$$\left( A_{2^L}^d f_s, \, (D_{2^j}^1 f_s)_{1 \le j \le L}, \, (D_{2^j}^2 f_s)_{1 \le j \le L} \right) .$$

A próxima etapa consiste em identificar um conjunto de feições pontuais nas subbandas LL das imagens de referência  $A_{2j}^d f_r$  (j=1,...,L) e de ajuste  $A_{2j}^d f_s$  (j=1,...,L), para cada nível da decomposição, usando o algoritmo descrito no Capítulo 4. As bandas  $D_{2j}^1 f_r$ ,  $D_{2j}^2 f_r$ ,  $D_{2j}^1 f_s$  e  $D_{2j}^2 f_s$  são usadas para estimar os gradientes das imagens de referência e de ajuste nas direções horizontal e vertical, para todos os níveis da decomposição, inclusive para o nível de resolução  $2^0$ .

#### 6.3 Transformação Espacial

As plataformas dos satélites, geralmente, têm condições de vôo e atitude estáveis, que fazem com que as imagens geradas sejam bem construídas. Neste caso, uma única função de transformação é suficiente para modelar a distorção das imagens (Buiten e Putten, 1997). No Capítulo 2 apresentamos algumas funções de transformação, que modelam diferentes distorções. A escolha adequada de uma destas funções é um dos principais problemas no processo de registro.

Porque os modelos polinomiais são simples e fáceis de serem usados, eles têm sido, muitas vezes, usados de forma incorreta e sem nenhum critério. A utilização dos modelos polinonomiais somente é válida se o grau do polinômio escolhido representar adequadamente as distorções geométricas das imagens.

Sistemas de informações geográficas, desenvolvidos mais recentemente, têm adotado o seguinte procedimento no processo de registro: 1) corrigem-se geometricamente as imagens usando as informações do sistema sensor; 2) registram-se as imagens usando os pontos de controle para refinar o registro (d'Alge, 1998). Após a primeira fase de correção, a geometria interna das imagens é satisfatória de forma que o uso de modelos simples são suficientes para refinar as correções. Esta etapa é realizada usando os pontos de controle.

Neste trabalho vamos considerar apenas as transformações afins e de similaridade. A extensão do método para o caso de polinômios de grau maior pode ser realizada e, portanto, isto não é colocado como uma limitação do nosso método.

Escolhida a função, os seus parâmetros são calculados usando a solução dos mínimos quadrados, o qual pode ser facilmente implementado e é razoavelmente rápido mesmo para um grande número de pontos (Duda e Hart, 1973).

No processo de reamostragem usamos a função de interpolação bilinear em todos os experimentos apresentados neste trabalho.

#### 6.4 Processo de Casamento

O processo de casamento de feições é realizado combinando técnicas de casamento baseadas em área e feições, discutidas no Capítulo 2. Dado o conjunto de feições pontuais detectadas nas sub-bandas LL das imagens de referência  $A_{2^j}^d f_r (j=1,...,L)$  e de ajuste  $A_{2^j}^d f_s (j=1,...,L)$ , o casamento pontual é realizado através da maximização do coeficiente de correlação (Equação 5.1),  $C_{f_r f_s}(X,Y,x,y)$ , sobre pequenas janelas  $(w_c x w_c)$  na vizinhança destes pontos. As coordenadas (X,Y) e (x,y) correspondem às posições das feições pontuais nas imagens de referência e de ajuste, respectivamente.

O uso do coeficiente de correlação como medida de similaridade é justificado pelo o seu bom desempenho, medido pela relação entre o número de casamentos corretos e alarmes falsos, o qual foi analisado no Capítulo 5. Além disso, a estratégia de multi-resolução permite reduzir o tempo de computação e a sensibilidade ao ruído nas imagens, que são as principais restrições ao uso dos métodos de correlação.

#### 6.4.1 Casamento Inicial

A Figura 6.2 mostra o diagrama de blocos das etapas envolvidas no processo de casamento inicial.

O casamento inicial é realizado nas sub-bandas LL no nível de resolução mais baixo da decomposição wavelet. Para cada ponto detectado na imagem de referência  $A_{2L}^d f_r$  procura-se um ponto na imagem de ajuste  $A_{2L}^d f_s$  que seja o seu mais similar de acordo com o critério de similaridade. Desde que  $A_{2L}^d f_r$  e  $A_{2L}^d f_s$  têm somente  $2^{-2L}N$  pixels, onde N é o número de pixels da imagem original, o processo de correlação sobre estas imagens é realizado com muito menos operações do que sobre as imagens originais. A correlação com uma janela quadrada de tamanho  $w_c$  sobre a imagem  $A_{2L}^d f_s$  é equivalente à correlação com uma janela de tamanho  $2^L w_c$  sobre a imagem original  $f_s$ , como mostra a Figura 6.3. Isto reduz o número de operações no processo de casamento quando realizado nos níveis de resolução mais baixos.

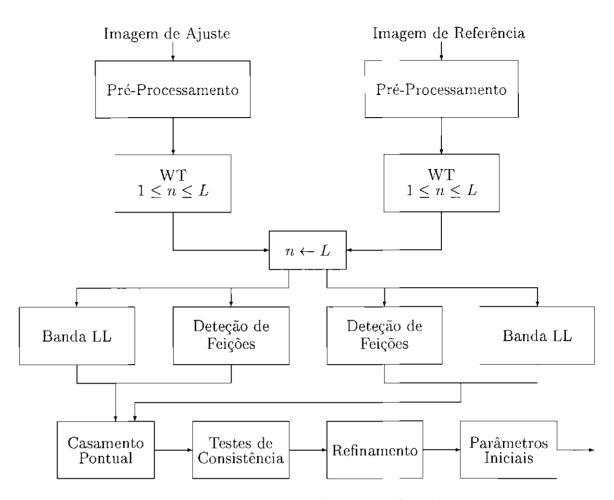


Fig. 6.2 - Processo de Casamento Inicial.

Sejam  $F_s = \{A^d_{2^L}f_s(x_i,y_i), i=1,\cdots,N_s\}$  e  $F_r = \{A^d_{2^L}f_r(X_j,Y_j), j=1,\cdots,N_r\}$  os conjuntos de feições detectadas para as imagens de ajuste e de referência na resolução  $2^L$ , respectivamente. Sejam  $C_{f_rf_s}$  e  $T_c$  o coeficiente de correlação e o valor do limiar acima do qual duas feições nas bandas  $A^d_{2^L}f_r$  e  $A^d_{2^L}f_s$  são consideradas similares. O ponto  $A^d_{2^L}f_r(X_k,Y_k)$  é dito o ponto mais similar ao ponto  $A^d_{2^L}f_s(x_t,y_t)$  se

$$C_{f_r f_s}(X_k, Y_k, x_l, y_l) = \max_{1 \le j \le N_s} C_{f_r f_s}(X_k, Y_k, x_j, y_j).$$
(6.1)

O processo de casamento é realizado da seguinte maneira. Para cada ponto  $A_{2L}^d f_r(X_j, Y_j) \in F_r$  todos os pontos  $A_{2L}^d f_s(x_j, y_j) \in F_s$  são examinados, e o ponto  $A_{2L}^d f_s(x_l, y_l)$  mais similar ao ponto  $A_{2L}^d f_r(X_k, Y_k)$  é selecionado.

Para assegurar a confiabilidade do casamento, testamos em seguida se a correlação

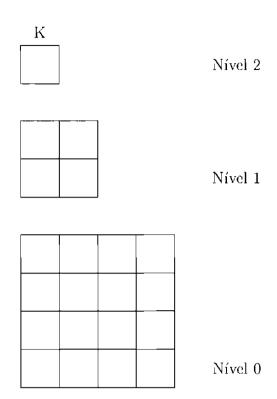


Fig. 6.3 - Uma janela de processamento de largura K na resolução baixa  $2^j$  corresponde às janelas de processamento de tamanho 2K e 4K nas resoluções mais finas  $2^{j-1}$  e  $2^{j-2}$ , respectivamente.

é razoavelmente alta. Se

$$C_{f,f_s}(X_k, Y_k, x_l, y_l) > T_c$$
, (6.2)

então  $A^d_{2^L}f_s(x_l,y_l)$  é chamado de o melhor casamento de  $A^d_{2^L}f_r(X_k,Y_k)$ .

# 6.4.2 Testes de Consistência

O processo de casamento inicial é a fase crucial em todo o processo de registro. Se os parâmetros de registro são estimados incorretamente a busca pela função de transformação poderá seguir na direção errada, não conseguindo mais recuperar o caminho correto nos passos posteriores. Para eliminar os pontos de controle suspeitos (alarmes falsos), dois testes de consistência dos casamentos são realizados:

1) Teste 1: verifica se o melhor casamento se dá nos dois sentidos (imagem de

referência ↔ imagem de ajuste) e

# 2) Teste 2: verifica se os casamentos são coerentes.

A seleção de uma feição  $A_{2L}^d f_s(x_l, y_l)$  na imagem de ajuste que melhor se casa com uma feição  $A_{2L}^d f_r(X_k, Y_k)$  na imagem de referência, não assegura que  $A_{2L}^d f_r(X_k, Y_k)$  seja o melhor casamento para  $A_{2L}^d f_s(x_l, y_l)$ , já que existem outros candidatos na imagem de referência que não foram testados. Então, devemos verificar se o melhor casamento para  $A_{2L}^d f_r(X_k, Y_k)$  existe e se é  $A_{2L}^d f_s(x_l, y_l)$ . Se este for o caso, ambos os pontos são casados. Ou seja, dado que  $A_{2L}^d f_s(x_l, y_l)$  é o candidato mais similar ao ponto  $A_{2L}^d f_r(X_k, Y_k)$ , o par de pontos  $[(X_i, Y_i), (x_i, y_i)]$  é casado se e somente se

$$C_{f_r f_s}(X_k, Y_k, x_l, y_l) > C_{f_r f_s}(X_j, Y_j, x_l, y_l)$$
,

para todo  $j \neq k$ .

Este casamento nos dois sentidos, denominado casamento bidirecional, reduz o número de falsos alarmes. Além disso, este casamento restritivo permite o uso de janelas de correlação menores. O esquema na Figura 6.4 ilustra o casamento bidirecional.

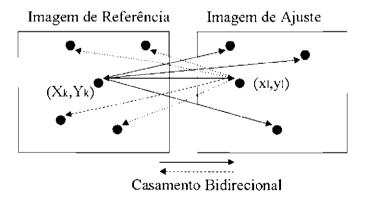


Fig. 6.4 - Casamento bilateral.

Após este primeiro teste alguns falsos casamentos ainda podem ocorrer, e devem ser eliminados para que os parâmetros da função de transformação sejam corretamente calculados.

Seja  $L_K = \{P_i = [(X_i, Y_i), (x_i, y_i)], i = 1, ..., K\}$  uma lista de K pares de pontos de controle, determinados no processo de casamento inicial. As coordenadas  $(X_i, Y_i)$  e  $(x_i, y_i)$  indicam as posições do iésimo par de pontos de controle nas imagens de referência e de ajuste, respectivamente.

Seja  $\mathcal{T}=(T_1,T_2)$  a estimativa da transformação que modela a distorção entre as imagens de referência e de ajuste. Para cada par de pontos de controle  $[(X_i,Y_i),(x_i,y_i)]\in L_K$  pode-se calcular o quanto cada ponto estimado pela transformação  $\mathcal{T},(T_1(x_i,y_i),T_2(x_i,y_i))$ , se afasta do ponto  $(X_i,Y_i)$  pelas seguintes expressões

$$rx_i = (T_1(x_i, y_i) - X_i)$$
  

$$ry_i = (T_2(x_i, y_i) - Y_i) ,$$
(6.3)

que são os chamados erros residuais nas direções x e y, respectivamente. O erro residual total  $r_i$  de cada ponto  $P_i$ ,

$$r_i = \sqrt{r_x^2 + r_y^2} \,, \tag{6.4}$$

mede o erro de ajuste do ponto ao modelo de mapeamento e pode ser usado para verificar a precisão do ponto de controle. Se o valor de  $r_i$  é um valor muito grande, pode-se inferir que o ponto  $P_i$  não é bom. No registro manual calculam-se os erros residuais para todos os pares de pontos de controle e eliminam-se ou ajustam-se aqueles cujos erros são altos. Similarmente, no registro automático o erro residual pode ser usado para eliminar os casamentos considerados falsos.

Para medir o erro global pode-se usar a tradicional medida raiz quadrada do erro médio quadrático (RMSE):

$$RMSE = \left(\frac{1}{N} \sum_{i=1}^{N} \left(rx_i^2 + ry_i^2\right)\right)^{1/2}.$$
 (6.5)

Valores altos de RMSE são indicativos da existência de casamentos errados ou com baixa precisão, se a deformação é modelada adequadamente e o número de pontos de controle é suficiente.

Quando a distorção entre as imagens é rígida, pode-se usar a hipótese de que as distâncias euclidianas entre dois pontos nas imagens de referência e de ajuste são preservadas. Neste caso, um procedimento similar ao de Li et al. (1995) e Welch et al. (1985) é usado para verificar a consistência dos casamentos e eliminar os falsos alarmes.

Sejam dois pares de pontos de controle  $P_k = [(X_k, Y_k), (x_k, y_k)]$  e  $P_l = [(X_l, Y_l), (x_l, y_l)] \in L_K$ . Sejam  $D_{lk}$  e  $d_{lk}$  as distâncias euclidianas entre os pares de pontos  $P_k$  e  $P_l$  nas imagens de referência e de ajuste, respectivamente:

$$D_{lk} = \sqrt{(X_l - X_k)^2 + (Y_l - Y_k)^2}$$

$$d_{lk} = \sqrt{(x_l - x_k)^2 + (y_l - y_k)^2},$$
(6.6)

Seja  $E_{ij}$  uma medida de erro que relaciona as distâncias Euclideanas  $D_{ij}$  e  $d_{ij}$  entre os pares de pontos de controle  $P_i$  e  $P_j$ :

$$E_{ij} = \frac{|D_{ij} - d_{ij}|}{\min\{D_{ij}, d_{ij}\}},\tag{6.7}$$

para i, j = 1, ..., K e  $i \neq j$ .

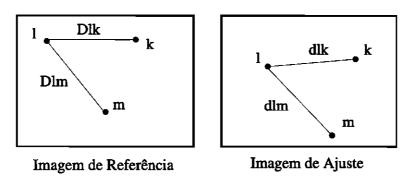


Fig. 6.5 - Teste de consistência: compara as distâncias entre todas as possíveis combinações de pares de pontos de controle.

Seja  $T_e$  um limiar de erro acima do qual um par de pontos de controle é considerado incorreto. Para uma transformação rígida, o procedimento de verificação

de consistência dos casamentos calcula e compara as distâncias euclidianas entre todas as possíveis K(K-1)/2 combinações de pares de pontos de controle (vide Figura 6.5). Os passos abaixo descrevem este procedimento:

- 1) Para cada dois pares de pontos de controle  $P_i$  e  $P_j$ , j = 1, ..., K, calculam-se as distâncias euclidianas  $D_{ij}$  e  $d_{ij}$  (Equação 6.6);
- 2) Para o ponto  $P_i$ , calculam-se os erros  $E_{ij}$  (Equação 6.7), para todo j = 1, ..., K;
- 3) Para cada  $P_i$  contam-se quantas vezes  $E_{ij} > T_e$ , para j = 1, ...K, e atribue-se este número a  $C_i$ , que é o contador de erros;
- 4) Repetem-se os passos 1-3 para todo  $P_i$ , i = 1, ..., K;
- 5) Remove-se da lista  $L_K$  o par de pontos de controle  $P_l$  cujo valor de  $C_l$  é o maior;
- 6) Atualizam-se os valores de  $C_k$ , para k = 1, ..., K e  $k \neq l$ ;
- 7) Atualiza-se o limiar  $T_e$ ;
- 8) Remove-se da lista  $L_K$  o próximo par mais incorreto;
- 9) Repete-se os passos 6 a 8 até que o erro médio quadrático dos pontos de controle (Equação 6.5) seja pequeno (< 1).

Este teste é realizado recursivamente de tal modo que os casamentos mais incorretos sejam removidos primeiro, seguidos pelos próximos mais incorretos, e assim sucessivamente. O valor do limiar  $T_e$  é reduzido gradativamente à medida que o erro RMSE torna-se pequeno

A idéia de relaxar a restrição do uso do algoritmo de verificação de consistência dos casamentos aos casos de imagens com distorção rígida fez com que pensássemos em outras soluções para o problema. O uso do procedimento de eliminação dos *outliers*, usado no registro manual e mencionado anteriormente, em algumas situações, não

funciona de forma satisfatória no caso do registro automático. No registro manual os pares de pontos de controle são escolhidos criteriosamente de forma que estes tenham correspondências corretas. O que pode acontecer é de alguns pares de pontos de controle não terem boa precisão e aí então serem eliminados ou corrigidos usando o procedimento de eliminação dos *outliers*. No caso do registro automático, tem-se um número total de pares de pontos de controle dentre os quais muitos são errados. Como estes também são usados para estimar os parâmetros do modelo de distorção, e se forem em grande número (por exemplo, maior do que o número de casamentos corretos), os pontos corretos acabam sendo descartados primeiro ficando somente os errados, invializando o uso do procedimento de eliminação dos *outliers*.

Para selecionar os casamentos corretos dentre os K casamentos obtidos na fase inicial, propomos um novo algoritmo que usa uma busca combinatória. Seja p o número mínimo de pontos de controle necessário para calcular os parâmetros da função de transformação escolhida e  $T_e$  um limiar de erro. O procedimento de eliminação dos alarmes falsos segue os seguintes passos:

- 1) Combinam-se os K pares de pontos de controle da lista  $L_K$ , em conjuntos de n = p + 1 elementos,  $C_i^n = \{P_1, P_2, ..., P_n\}, i = 1, ..., \frac{K!}{n!(K-n)!};$
- 2) Para cada conjunto  $C_i^n$ , calcula-se o erro (Equação 6.5),  $E_i^n$ , usando os pares de pontos de controle do conjunto;
- 3) Seleciona-se o conjunto  $C_k^n$  com o menor erro  $E_k^n$ ;
- 4) Para todo  $P_j \in L_K$  e  $P_j \not\in C_k^n, \ j=1,...,K$  faz-se:
  - adiciona-se um par de pontos  $P_j$  ao conjunto  $C_k^n o C_{kj}^n$ ;
  - calcula-se o erro do conjunto  $C_{kj}^n$ ,  $E_{kj}^n$ ;
- 5) Seleciona-se o conjunto  $C_{kj}^n$ , dentre todos os gerados no passo anterior, com o menor erro  $E_{kj}^n$ ;
- 6) Se  $E_{kj}^n > T_e$ , para-se;

- 7) Faz-se  $C_k^n = C_{kj}^n$ ;
- 8) Repetem-se os passos 4 a 7;

A possibilidade dos p+1 elementos do conjunto inicial serem errados e mesmo assim apresentarem um erro inicial pequeno, existe. Mas, neste caso, à medida que se adicionam pontos o erro tende aumentar rapidamente, e podemos detectar a situação desfavorável, interrompendo o processo, e reiniciá-lo, escolhendo o próximo conjunto com menor erro. Observa-se que, para todos os casos testados neste trabalho, o método teve sucesso.

### 6.4.3 Refinamento dos pontos de controle iniciais

Após os casamentos passarem pelos testes de consistência temos no final uma lista de N pares de pontos de controle  $L_N = \{[(X_i, Y_i), (x_i, y_i)], i = 1, 2, ..., N\}$ . Devido ao processo de filtragem e dizimação há uma certa imprecisão nas posições das feições extraídas, que acarretam por sua vez uma imprecisão nos pontos de controle. Esta imprecisão é corrigida à medida que o processamento é realizado nos níveis de resolução mais fina.

No nível de resolução mais baixa o refinamento é realizado da seguinte maneira. Para cada par de pontos de controle  $[(X_i, Y_i), (x_i, y_i)] \in L_N$ ,  $(X_i, Y_i)$  é casado com o ponto  $(x_l, y_l)$  se:

$$C_{f_s f_r}(X_i, Y_i, x_l, y_l) = \max_{-w_s/2 \le m, n \le w_s/2} C_{f_s f_r}(X_i, Y_i, x_i + m, y_i + n), \tag{6.8}$$

onde  $w_s$  (inteiro e ímpar) é o tamanho da janela de refinamento. Ou seja, para um ponto na posição  $(x_l, y_l)$  na imagem de ajuste procura-se na vizinhaça do ponto na posição  $(X_l, Y_l)$  na imagem de referência aquele cujo valor da correlação é máximo.

Este refinamento tem como objetivo melhorar a precisão dos pontos de controle no nível de resolução mais baixa. Os pares de pontos de controle obtidos após esta fase de refinamento são usados para calcular a função de mapeamento inicial.

# 6.4.4 Algoritmo de Casamento Inicial

Seja L o número de níveis de resolução da decomposição wavelet. Combinando as operações descritas na seção anterior, descrevemos de forma resumida as etapas envolvidas no processo de casamento de feições nas imagens de referência  $A^d_{2^j}f_r$  e de ajuste  $A^d_{2^j}f_s$ , no nível de resolução L.

O processo de casamento no nível L é realizado da seguinte maneira:

- 1) Calcula-se a decomposição wavelet;
- Extraem-se as feições pontuais mais significativas no nível L, para as imagens de referência e de ajuste;
- 3) Eliminam-se os pontos de feições próximas às bordas das imagens;
- 4) Realiza-se o casamento bidirecional de feições usando o coeficiente de correlação  $C_{f_rf_s}(X,Y,x,y)$ ;
- 5) Verifica-se se os casamentos obtidos são coerentes;
- 6) Refinam-se os pontos de controle (Equação 6.8);
- 7) Determinam-se os parâmetros da função de transformação inicial  $\mathcal{T}_L$ ;
- 8) Vai-se para o próximo nível L-1

#### 6.5 Processo de Refinamento do registro

No algoritmo de refinamento os parâmetros da função de transformação, obtidos na resolução  $2^{j}$ , são usados para transformar a imagem de ajuste no próximo nível de resolução mais fina  $2^{j-1}$ . A Figura 6.6 mostra o diagrama de blocos das etapas envolvidas no processo de refinamento proposto.

No processo de refinamento o resultado do processo de casamento na resolução mais baixa  $2^j$  é usado para restringir a região de busca no processo de casamento na

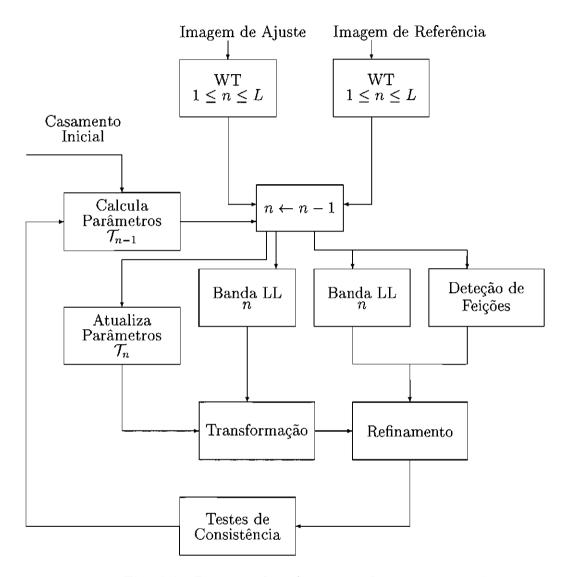


Fig. 6.6 - Processo de refinamento dos parâmetros de  ${\rm transformação}.$ 

próxima resolução mais fina  $2^{j-1}$ . A região ao redor de um ponto  $(x^j, y^j)$  na imagem  $A^d_{2^j} f_s$  corresponde na imagem  $A^d_{2^{j-1}} f_s$  a uma região ao redor de um dos pontos

$$\{(2x^j, 2y^j), (2x^j + 1, 2y^j)(2x^j, 2y^j + 1), (2x^j + 1, 2y^j + 1)\}$$
.

Similarmente,  $(X^j,Y^j)$  em  $A^d_{2^j}f_r$  corresponde a uma posição próxima ao ponto  $(2X^j,2Y^j)$  em  $A^d_{2^{j-1}}f_r$ . Dado um par de pontos casados  $[(X^j,Y^j),(x^j,y^j)]$  na resolução  $2^j$ , para procurar o ponto  $(x^{j-1},y^{j-1})$  na imagem  $A^d_{2^{j-1}}f_s$  correspondente ao ponto  $(X^{j-1},Y^{j-1})$  em  $A^d_{2^{j-1}}f_r$  basta calcular a medida de similaridade na vizinhança de  $(x^{j-1},y^{j-1})=(2x^j,2y^j)$  em  $A^d_{2^{j-1}}f_s$  restringindo o espaço de busca.

Seja  $A^d_{2^{j-1}}f^t_s$  a banda LL registrada usando os parâmetros da função de transformação obtidos no nível de resolução  $2^j$  e atualizados para o nível  $2^{j-1}$ . O processo de casamento na fase de refinamento é realizado usando  $A^d_{2^{j-1}}f^t_s$  e o conjunto de pontos detectados na imagem de referência na resolução  $2^{j-1}$ . Cada ponto  $(X^{j-1}_k, Y^{j-1}_k)$  em  $A^d_{2^{j-1}}f_r$  é casado com o ponto  $(x^{j-1}_l, y^{j-1}_l)$  em  $A^d_{2^{j-1}}f_s$  se e somente se

$$\begin{split} C_{f_s^tf_r}(X_k^{j-1}, & Y_k^{j-1}, x_l^{j-1}, y_l^{j-1}) \ = \ \max_{-w_s/2 \le m, n \le w_s/2} C_{f_s^tf_r}(X_k^{j-1}, Y_k^{j-1}, X_k^{j-1} + m, Y_k^{j-1} + n), \\ C_{f_s^tf_r}(X_k^{j-1}, Y_k^{j-1}, x_l^{j-1}, y_l^{j-1}) \ > \ 0, 8 \ , \end{split}$$

onde  $w_s$  (inteiro e ímpar) é o tamanho da janela de refinamento. Uma janela de tamanho  $w_s = 5$  é suficiente para cobrir a região no nível de resolução  $2^{j-1}$  projetada por uma região no nível de resolução  $2^j$ . Em cada nível de refinamento, os casamentos obtidos são testados. Seja  $L_K$  uma lista de K pares de pontos de controle. Aqueles pares de pontos com erro residual de registro altos são sequencialmente eliminados de modo iterativo, usando o seguinte procedimento:

- 1) Para cada par de pontos de controle da lista  $L_K$ ,  $P_i$ , calcula-se o erro  $r_i$  (Equação 6.4), para todo j=1,...,K;
- 2) calcula-se o valor de RMSE para a lista de pontos de controle obtida;
- 3) Se  $RMSE \ge T_e$ , remove-se da lista o par de pontos de controle com o maior erro residual  $r_i$ ;

- 4) Recalcula-se o valor de RMSE para os pontos de controle restantes;
- 5) Atualiza-se o limiar  $T_e$ ;
- 6) Se  $RMSE < T_e$  pára-se;
- 7) Remove-se da lista o próximo par mais incorreto;
- 8) Repetem-se os passos 2 a 6 até que o erro RMSE seja pequeno (< 1).

Os pontos de controle refinados são usados para calcular os parâmetros de refinamento  $\{s_{j-1}^r, \theta_{j-1}^r, \Delta x_{j-1}^r, \Delta y_{j-1}^r\}$  da transformação de similaridade  $\mathcal{T}_{j-1}^r$ . Os parâmetros da transformação total  $\mathcal{T}_{j-1}$ ,  $\{s_{j-1}, \theta_{j-1}, \Delta x_{j-1}, \Delta y_{j-1}\}$ , são obtidos pela combinação das transformações  $\mathcal{T}_{j-1}^r$  e  $\mathcal{T}_j$  (Apêndice C),

$$\begin{pmatrix} s_{j-1} \\ \theta_{j-1} \\ \Delta x_{j-1} \\ \Delta y_{j-1} \end{pmatrix} = \begin{pmatrix} s_{j-1}^r s_j \\ \theta_{j-1}^r + \theta_j \\ 2s_{j-1}^r (\cos(\theta_{j-1}^r) \Delta x_j + \sin(\theta_{j-1}^r) \Delta y_j) + \Delta x_{j-1}^r \\ 2s_{j-1}^r (\cos(\theta_{j-1}^r) \Delta x_j + \sin(\theta_{j-1}^r) \Delta y_j) + \Delta y_{j-1}^r \end{pmatrix}.$$
(6.9)

No caso da transformação afim genérica, os parâmetros no nível de resolução  $2^{j-1}$ ,  $\{a_{1j-1}, a_{2j-1}, a_{3j-1}, b_{1j-1}, b_{2j-1}, b_{3j-1}\}$ , são obtidos pela combinação das transformações  $\mathcal{T}_{j-1}^r$  e  $\mathcal{T}_j$ :

$$\begin{pmatrix}
a_{1j-1} \\
a_{2j-1} \\
a_{3j-1} \\
b_{1j-1} \\
b_{2j-1} \\
b_{3j-1}
\end{pmatrix} = \begin{pmatrix}
a_{1j-1}^{r} a_{1j} + a_{2j-1}^{r} b_{1j} \\
a_{1j-1}^{r} a_{2j} + a_{2j-1}^{r} b_{2j} \\
a_{1j-1}^{r} a_{3j} + a_{2j-1}^{r} b_{3j} + a_{3j} \\
b_{1j-1}^{r} a_{1j} + b_{2j-1}^{r} b_{1j} \\
b_{1j-1}^{r} a_{2j} + b_{2j-1}^{r} b_{2j} \\
b_{1j-1}^{r} a_{3j} + b_{2j-1}^{r} b_{3j}
\end{pmatrix}.$$
(6.10)

Em cada nível da pirâmide os parâmetros da função de mapeamento são atualizados usando os parâmetros obtidos na fase de refinamento e os obtidos no nível anterior.

Esta operação é realizada de forma similar nas resoluções progressivamente maiores. Depois de processar todos os níveis, ainda resta processar as imagens originais (resolução 2º). O mesmo procedimento de casamento e refinamento dos pontos de controle, descrito anteriormente é aplicado para obter a função de transformação final usada para registrar a imagem de ajuste original.

No último nível de processamento, um conjunto de pontos de controle mais precisos do que os da fase inicial, é obtido. Consequentemente, um registro mais preciso é obtido.

# 6.5.1 Algoritmo de Refinamento

Seja L o número de níveis de resolução da decomposição wavelet. Combinando as operações descritas nas seções anteriores, resumimos aqui as etapas envolvidas no processo de refinamento do casamento de feições nas imagens de referência  $A^d_{2j} f_r$  e de ajuste  $A^d_{2j} f_s$ , nos diferentes níveis de resolução, l=0,1,...,L.

Seja  $\mathcal{T}_j$  a função de transformação obtida no nível j=L, na fase de casamento inicial. O processo de refinamento é realizado da seguinte maneira:

- 1) Vai-se para o próximo nível da decomposição j-1
- 2) Extraem-se as feições pontuais mais significativas;
- 3) Eliminam-se os pontos de feições próximas às bordas das imagens;
- 4) Atualiza-se a função de transformação no nível  $j-1, \mathcal{T}_{j-1}$ .
- 5) Registra-se a imagem de ajuste  $(A_{2^{j-1}}^d f_s \to A_{2^{j-1}}^d f_s^t)$ , usando  $\mathcal{T}_{j-1}$ .
- 6) Usa-se a estimativa  $\mathcal{T}_{j-1}$  para limitar a região de busca;
- 7) Faz-se o refinamento dos casamentos usando as sub-imagens  $A_{2^{j-1}}^d f_s^t$ , e  $A_{2^{j-1}}^d f_r$  e as feições extraídas no nível j-1 dentro da região de busca;
- 8) Determina-se a função de transformação residual,  $\mathcal{T}_{j-1}^r$ .

- 9) Atualizam-se os parâmetros da função de transformação no nível j-1, considerando os parâmetros de refinamento (Equações 6.9 e 6.10);
- 10) Se o nível corrente é o nível 0 (imagens originais):
  - repetem-se os passos 2 a 5;
  - Usa-se a estimativa  $\mathcal{T}_0$  para limitar a região de busca;
  - faz-se o refinamento dos casamentos usando as imagens  $f_s^t$  e  $f_r$  e as feições extraídas no nível de resolução  $2^0$  dentro da região de busca;
  - Atualizam-se os parâmetros da função de mapeamento  $\mathcal{T}_0$  após o refinamento (Equações 6.9 e 6.10);
  - Registra-se a imagem de ajuste  $f_s$  e para-se;
- 11) Retorna-se para o passo 1.

No processo de refinamento, a operação de casamento dos pontos é realizada somente para aqueles pontos que caem dentro da área de sobreposição entre as imagens de referência e de ajuste após o registro. Isto elimina processamentos desnecessários tornando o processamento mais rápido.

Alguns resultados preliminares são apresentados em Fonseca e Costa (1997) e Fonseca et al.(1998). O algoritmo é aplicado no registro de imagens multi-temporais, adquiridas em áreas de agricultura, floresta e urbana, típicas da área de sensoriamento remoto. Os resultados do registro foram satisfatórios, com uma precisão de registro menor do que um pixel.

# CAPÍTULO 7

#### RESULTADOS EXPERIMENTAIS

Para a utilização do algoritmo de registro proposto, alguns parâmetros devem ser definidos pelo usuário. Os principais parâmetros são o número de níveis da decomposição wavelet L e o valor de  $\beta$  (Equação 4.7). A escolha do valor do parâmetro L depende do tamanho das imagens a serem processadas. Quando as sub-imagens, no nível de resolução mais baixa, são muito pequenas (por exemplo, menores que  $64 \times 64$  pixels) a dificuldade para se obter um número suficiente de pontos de controle para realizar o registro aumenta. O valor de  $\beta$  define o número de feições extraídas nas imagens. Como o processo de casamento de feições é a que demanda mais tempo, um número de feições muito grande torna muito lento o processamento. Por outro lado, um número de feições pequeno pode não ser suficiente para gerar uma estimativa correta dos parâmetros da função de transformação.

Valores típicos dos parâmetros do algoritmo usados nos experimentos apresentados aqui são:  $w_c = 11$  ou  $wc_{\pm}13$  (Equação 5.1)),  $T_c = 0,75$  (Equação 6.2) e  $T_w = 0,9$  (Equação 4.10). Quando as imagens possuem variações sazonais significativas entre si, o valor de  $w_c$  pode ser modificado para levar em conta o baixo grau de similariade entre as imagens.

A seguir vários resultados experimentais são realizados para ilustrar o desempenho do método proposto.

## 7.1 Experimento A

Para verificar se o método é preciso e robusto algumas imagens são distorcidas artificialmente e corrigidas usando o método. Para este experimento selecionamos seis imagens de sensores e regiões diferentes. As Figuras 7.1(a)-(f) mostram as imagens originais usadas para gerar as imagens simuladas usando os parâmetros do modelo de distorção da Tabela 7.1 (real).

**TABELA 7.1 -**PARÂMETROS DO MODELO DE DISTORÇÃO REAIS E ESTIMADOS PELO MÉTODO DE REGISTRO PROPOSTO

Imagens	área	Parâmetros	S	θ	$\Delta x$	$\Delta y$
AES-1	agric./floresta	real	1,1	9,0	100,0	70,0
(radar)	(600×600)	estimado	1,1	8,99	99,68	69,97
fotografia	urbana	real	1,0	-12,0	-100,0	-50,0
aérea	(600×600)	estimado	1,0	-12,0	-100,1	-49,92
SIR-C	agricultura	real	0,92	8,0	80,0	-20,0
(radar)	(512×512)	estimado	0,919	8,0	79,97	-19,92
sensor óptico	deserto	real	0,90	15,0	38,0	-55,0
(balão)	(512×512)	estimado	0,90	14,99	37,98	-54,97
Landsat-TM	floresta	real	1,10	20,0	45,0	-100,0
	(1000×1000)	estimado	1,10	20,01	45,12	-100,10
SPOT	mata/cerrado	real	0,95	10,30	385,0	201,0
	(1000×1000)	estimado	0,95	10,29	384,56	200,63

A Figura 7.1(a) é uma imagem de radar aerotransportado AES-1 (Empresa Aerosensing Radarsystems, Alemanha), de uma área de agricultura e floresta de pinus, na região de Wessling, Alemanha, banda X, 600×600 pixels, reamostrada para um pixel de tamanho 2 metros (resolução original 0,5 m.). A fotografia aérea da cidade de São José dos Campos, 600×600 pixels, tomada em 08/04/76 é mostrada na Figura 7.1(b). Uma imagem de radar adquirida em 13/04/94 através da missão do space shuttle SIR-C/X-SAR (Spaceborne Imaging Radar C/X-Band Synthetic Aperture Radar) na região de Bebedouro, Pernambuco, banda L, polarização HH, pixel = 12,5 metros, 512×512 pixels, é mostrada na Figura 7.1(c). Uma imagem adquirida por um sensor ótico acoplado a um balão, de uma região do deserto de Mojave, Califórnia, é mostrada na Figura 7.1(d). A Figura 7.1(e) mostra a imagem Landsat-TM, banda 3, 1000×1000 pixels, adquirida em 08/1991 na região de Manaus, Amazônia. A Figura 7.1(f) mostra uma imagem da cidade de Brasília, Distrito Federal, com uma área de mata ciliar (mata ao longo de drenagens) e cerrado, adquirida em 1989 pelo satétile SPOT, banda pancromática (resolução de 10 m.), 1000×1000 pixels.

Cada uma das imagens testes é transformada usando o modelo de similaridade com os seguintes parâmetros: escala (s), rotação  $(\theta)$  e deslocamento em x  $(\Delta x)$  e y  $(\Delta y)$ . A Tabela 7.1 mostra os tipos de dados usados (colunas 1 e 2), os parâmetros do modelo de distorção usado para simular as imagens (real) e os respectivos parâmetros estimados pelo método proposto (estimado). Nesta tabela, as unidades de medidas para  $\theta$  é em graus e para  $\Delta x$  e  $\Delta y$  é em pixels. Pode-se observar, pela tabela, que os parâmetros são estimados com boa precisão em todos os casos, até mesmo no experimento 5 (imagem Landsat-TM) em que a rotação é relativamente grande  $(\theta=20)$ . Os parâmetros do algoritmo,  $w_c=13$ ,  $T_c=0$ , 75 e  $T_w=0$ , 9, são iguais para todos os casos. Os valores de  $\beta$  são ajustados para gerar um número de feições suficiente para estimar a função de transformação inicial e assim realizar o registro com sucesso. As imagens maiores  $(1000\times1000)$  foram processadas em L=3 níveis, e as outras em L=2 níveis de resolução.

Para ilustrar o desempenho do nosso algoritmo, apresentamos a seguir alguns resul-

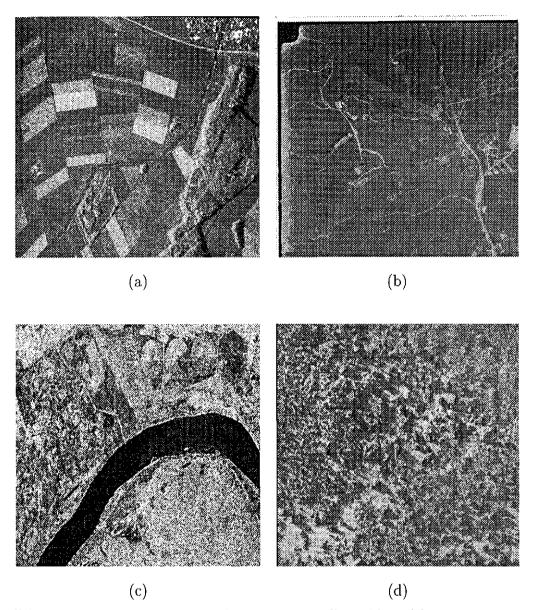


Fig. 7.1 - Imagens usadas para simular as imagens distorcidas: (a) imagem de radar AES-1, banda X (600×600), região de agricultura/floresta; (b) imagem de uma foto aérea da cidade de São José dos Campos (600×600); (c) imagem de radar do SIR-C/X, banda L, de uma área de agricultura (512×512); (d) imagem do deserto Mojave, adquirida por um sensor ótico acoplado à um balão (512×512); (e) imagem Landsat-TM de uma região de floresta em Manaus, AM (1000×1000) e (f) imagem do satélite SPOT, banda pancromática (10 metros), de uma região da cidade de Brasília, DF (1000×1000).

(continua)

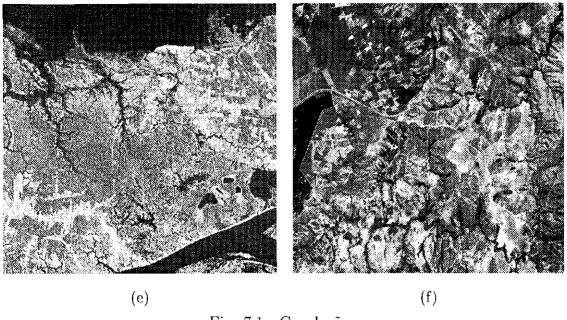


Fig. 7.1 - Conclusão

tados obtidas nas etapas de simulação e correção das imagens. As Figuras 7.2, 7.3, 7.4 e 7.5(a) e (b), mostram os pontos de controle iniciais sobrepostos às bandas LL das imagens original e simulada, com os parâmetros da Tabela 7.1, para as imagens das Figuras 7.1(a), (c), (d) e (e), respectivamente.

As Figuras 7.6(a), (b), (c) e 7.7(a), (b) e (c) mostram a imagem original, a imagem simulada e o mosaico após o registro, respectivamente, para as imagens AES-1 e Landsat-TM. No caso das imagens Landsat-TM, imagens menores (400×400) foram extraídas das imagens original e simulada (1000×1000) e a seguir registradas.

# 7.2 Experimento B

Para verificar a precisão do algoritmo no registro de duas imagens que apresentam diferenças sazonais, cujo modelo de distorção das imagens é conhecido, realizamos o seguinte experimento. Registramos duas imagens da região amazônica, adquiridas em épocas diferentes, TM95 (1995) e TM97 (1997)), banda 5, de tamanhos  $1000\times1000$  pixels. As imagens foram registradas manualmente, com precisão de registro de 0,52 pixel. Uma das imagens (TM97) foi transformada pela função de similaridade com os seguintes parâmetros:  $s=0,90; \theta=10,0; \Delta x=-300,0$  e

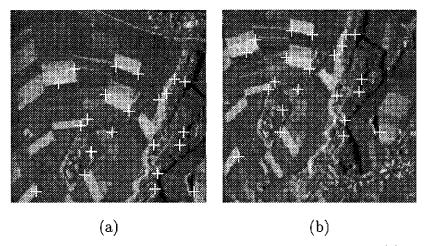


Fig. 7.2 - Pontos de controle, sobrepostos às bandas LL das imagens (a) original e (b) simulada, obtidos para corrigir a imagem EAS-1 distorcida (L=2,  $\beta$  = 1).

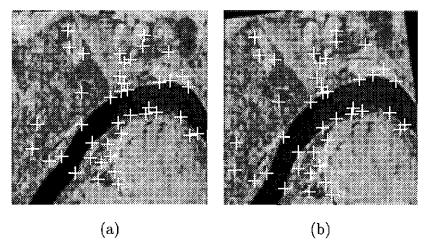


Fig. 7.3 - Pontos de controle, sobrepostos às bandas LL das imagens (a) original e (b) simulada, obtidos para corrigir a imagem SIR-C/X distorcida (L=2,  $\beta$  = 2).

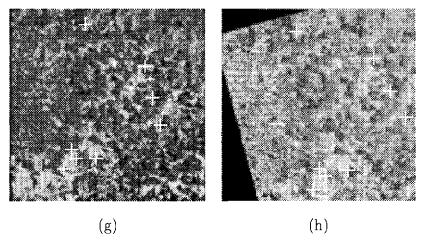


Fig. 7.4 - Pontos de controle, sobrepostos às bandas LL das imagens (a) original (sensor ótico acoplado ao balão) e (b) simulada, obtidos para corrigir a imagem distorcida (L=2,  $\beta=1$ ).

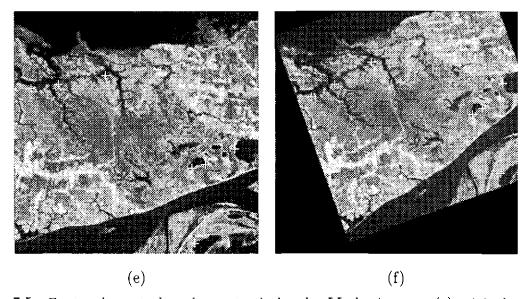


Fig. 7.5 - Pontos de controle, sobrepostos às bandas LL das imagens (a) original e (b) simulada, obtidos para corrigir a imagem Landsat-TM distorcida (L=3,  $\beta$  = 2).

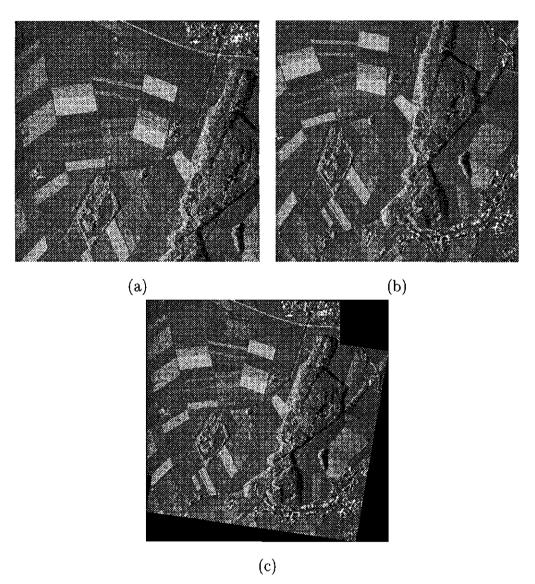


Fig. 7.6 - Mosaico entre as imagens do sistema AES-1 original e simulada corrigida: (a) imagem de referência: imagem AES-1 original; (b) imagem de ajuste: imagem simulada; (c) mosaico, após o registro (L=2,  $\beta=1$ ).

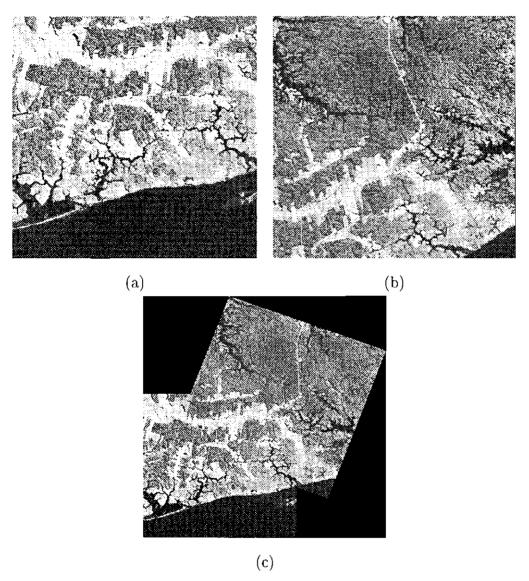


Fig. 7.7 - Mosaico entre as imagens Landsat-TM e a simulada, após a correção: (a) imagem de referência: original ( $400\times400$  pixels); (b) imagem de ajuste: imagem simulada ( $400\times400$  pixels); (c) mosaico, após o registro (L=2,  $\beta=3$ ).

 $\Delta y = -350, 0$ . Em seguida, a imagem distorcida foi registrada com a imagem original, usando o método de registro proposto. Os parâmetros do modelo de distorção foram corretamente estimados com uma boa precisão, a saber:  $s = 0,90; \theta = 9,976; \Delta x = -299,76$  e  $\Delta y = -350,46$ . Para ilustrar o resultado, as Figuras 7.8(a) e (b) mostram os pontos de controle sobrepostos as bandas LL das imagens de referência TM95 e TM97 simulada, obtidos no casamento inicial. A Figura 7.9 mostra a composição colorida das imagens registradas para ilustrar como a imagem corrigida se sobrepõe perfeitamente à imagem de referência TM95. Pode-se observar que, embora a área de sobreposição entre as imagens multi-temporais seja razoavelmente pequena, o registro foi realizado com sucesso.

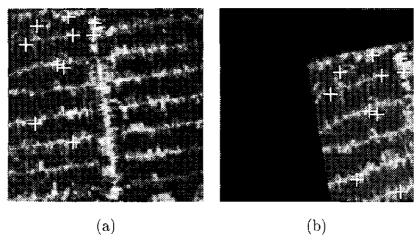


Fig. 7.8 - Pontos de controle, sobrepostos às bandas LL das imagens (a) original (TM95) e (b) TM97 distorcida (L=3,  $\beta = 2$ ).

## 7.3 Experimento C

Um importante projeto no INPE é o estudo da dinâmica do uso e ocupação do solo através de imagens multi-temporais para o monitoramento da região Amazônica (Alves, 1992; Batista et al., 1994). Uma das etapas do projeto trata-se do registro das imagens, o qual é realizado manualmente com um alto custo operacional. A automatização do processo de registro tornaria mais rápido esta tarefa. Devido à característica textural forte destas imagens e ao fato de serem multi-temporais, os métodos de registro baseados em contornos fechados, como o proposto por Li et



Fig. 7.9 - Composição colorida das imagens TM97 corrigida e TM95: canal R - TM97 corrigida, canal G - TM95, canal B - TM95.

al. (1995), não funcionam satisfatoriamente. Para mostrar a eficiência do nosso método no registro dessa classe de imagens, selecionamos três pares de imagens óticas multi-temporais da região Amazônica.

O algoritmo proposto é também aplicado para registrar imagens de regiões de agricultura e áreas urbanas. Imagens dos satélites SPOT, Landsat-TM são usados neste experimento. As informações das imagens testes usadas nos experimentos são descritas na Tabela 7.2. Os pares de imagens testes, são de tamanho 512×512 e foram extraídas de imagens maiores afim de reduzir o custo computacional e espaço em disco.

As Figuras 7.10-7.17 apresentam os resultados do registro para os pares de imagens testes da Tabela 7.2. Para a maioria dos testes (imagens  $512\times512$ ) as imagens foram processadas usando a decomposição wavelet em L=2 níveis e os parâmetros  $\beta=3, w_c=13, T_c=0.75$  e  $T_w=0.9$ . O modelo de distorção usado, em todos os testes apresentados, é o da similaridade. As imagens Landsat-TM e SPOT, usadas nos experimentos, possuem uma boa qualidade geométrica e, portanto, o modelo de similaridade é suficiente para modelar as suas distorções.

Α **Figura** 7.10 resultado registro das mostra do imagens TM925AM e TM945AM. Nosso algoritmo de registro foi aplicado, usando a imagem TM945AM (Figura 7.10(a)) como referência e a imagem TM925AM (Figura 7.10(b)) como a de ajuste. As Figuras 7.10(c), (d) e (e) mostram os pontos de controle obtidos no casamento inicial e a imagem TM925AM após a transformação, respectivamente. Os pontos de controle (marcados com sinais de +) são sobrepostos às bandas LL do nível mais baixo da decomposição wavelet, que são ampliadas para melhor visualização dos pontos de controle.

No segundo teste a imagem TM955AM (Figura 7.11(a)) foi tomada como imagem de referência e a imagem TM975AM (Figura 7.11(b)) como imagem de ajuste. Os pontos de controle iniciais e a imagem TM97F, após o registro, são mostrados nas Figuras 7.11(c), (d) e (e), respectivamente. Do mesmo modo, a Figura 7.12 mostra os

TABELA 7.2 - IMAGENS TESTES (ÓTICAS)

Imagens	região	satélite	banda	data	área
TM925AM	Amazônia	Landsat-TM	5	07/06/92	floresta
TM945AM		Landsat-TM	5	15/07/94	
TM975AM	Amazônia	Landsat-TM	5	07/07/97	floresta
TM955AM		Landsat-TM	5	03/08/95	
TM965AM	Amazônia	Landsat-TM	5	20/07/96	floresta
TM945AM		Landsat-TM	5	15/07/94	
TM905IT	Itapeva	Landsat-TM	5	09/09/90	agricultura
TM945IT		Landsat-TM	5	18/07/94	
TM905IT	Itapeva	Landsat-TM	5	09/09/90	agricultura
TM945RIT		Landsat-TM	5	18/07/94	(georef.)
TM945AG	Agudos	Landsat-TM	5	09/07/94	agricultura
TM925AG		Landsat-TM	5	21/09/92	
SP953SP	São Paulo	SPOT	3	08/08/95	urbana
TM944SP		Landsat-TM	4	07/06/94	
SP953DF	Brasília	SPOT	3	08/08/95	urbana
TM944DF		Landsat-TM	4	07/06/94	

resultados do registro das imagens TM945AM e TM965AM. A imagem TM945AM foi tomada como referência e a imagem TM965AM como imagem de ajuste.

No quarto experimento a imagem TM945I (Figura 7.13(a)) foi escolhida como a imagem de referência e a imagem TM905I (Figura 7.13(b)) como a de ajuste. A imagem TM905I, após a transformação, é mostrada na Figura 7.13(e). As Figuras 7.13(b) e (c) mostram os pontos de controle obtidos no processo de casamento inicial. A imagem mostrada na Figura 7.14(b) é o resultado do registro da imagem de referência TM945RI na Figura 7.14(a) e a imagem de ajuste TM905I na Figura 7.13(b). A imagem TM945RI, tomada como referência, foi antes registrada manualmente com um mapa e por isso apresenta uma pequena rotação em relação a imagem TM945I da Figura 7.13(a). As Figuras 7.14(c) e (d) mostram os pontos de controle para este experimento.

Imagens de uma área de agricultura, próxima a cidade de Agudos, São Paulo, adquiridas em um intervalo de tempo de dois anos, foram registradas. As Figuras 7.15(a) e (b) mostram as imagens de referência (TM925AG) e de ajuste (TM945AG). Os pontos de controle iniciais e o resultado do registro são apresentados nas Figuras 7.15(c), (d) e (e).

Os dois experimentos a seguir, foram realizados para registrar dois pares de imagens da área urbana da cidade de São Paulo e de Brasília, DF. Uma imagem SPOT da cidade de São Paulo (SP953SP), banda 3, foi reduzida para um tamanho de pixel de 30 m e selecionada como a imagem de ajuste (Figura 7.16(b)). Outra imagem Landsat-TM, banda 4 da mesma região (TM944SP) foi selecionada como imagem de ajuste (Figura 7.16(b)). As bandas 3  $(0,79-0,84\mu m)$  do SPOT e 4  $(0,76-0,90\mu m)$  do TM ocupam, praticamente, a mesma faixa do espectro eletromagnético e possuem uma boa resposta a alvos de área urbana. Por isso foram selecionadas. Os pontos de controle e o resultado do registro são apresentados nas Figuras 7.16(c), (d) e (e), respectivamente. Da mesma forma, a imagem da cidade de Brasília adquirida pelo satélite SPOT (SP953DF) foi reamostrada e tomada como imagem de ajuste. A imagem TM (TM944DF) foi selecionada como a imagem de referência. As Fi-

guras 7.17(a) e (b) mostram as imagens originais e as Figuras 7.17(c), (d) e (e) mostram os pontos de controle iniciais e o resultado do registro. A redução de escala das imagens SPOT (fator 0,67), nos dois casos, foi realizada através de uma interpolação bilinear.

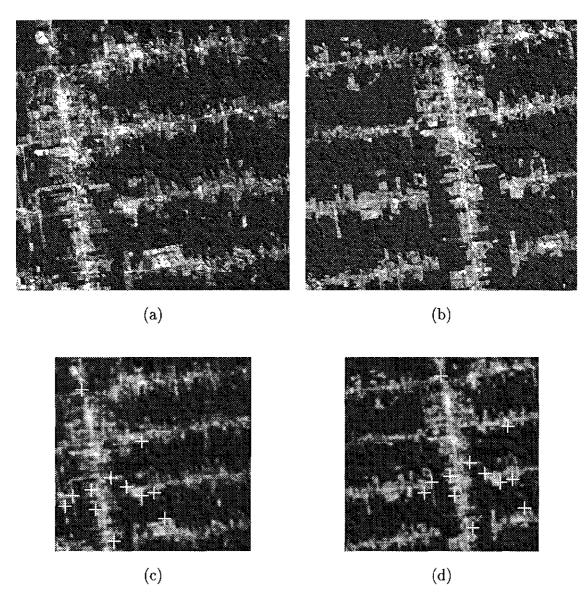


Fig. 7.10 - Imagens da Amazônia: (a) imagem de referência - TM945AM; (b) imagem de ajuste - TM925AM; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM945AM e TM925AM, respectivamente; (e) imagem TM925AM após a transformação final (L=2 e  $\beta=3$ ).

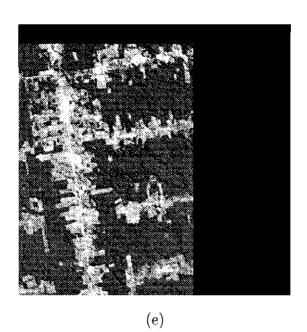


Fig. 7.10 - Conclusão

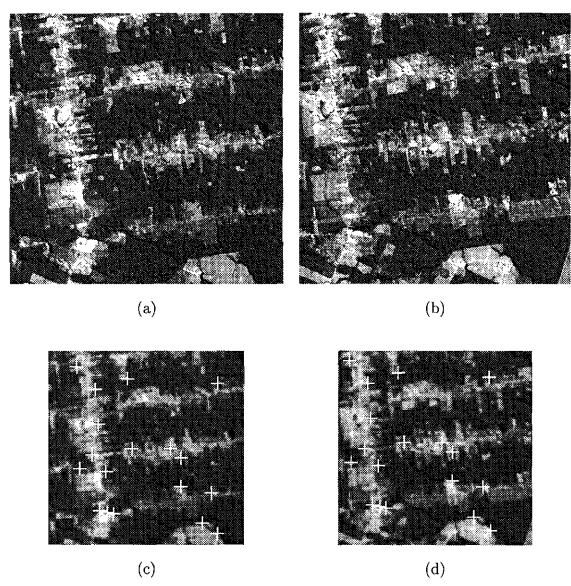


Fig. 7.11 - Imagens da Amazônia: (a) imagem referência - TM955AM; (b) imagem de ajuste - TM975AM; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM955AM e TM975AM, respectivamente; (e) Imagem TM975AM após a transformação final (L=2 e  $\beta=3$ ).

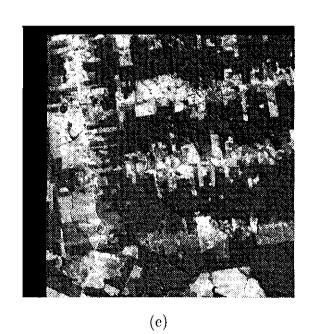


Fig. 7.11 - Conclusão

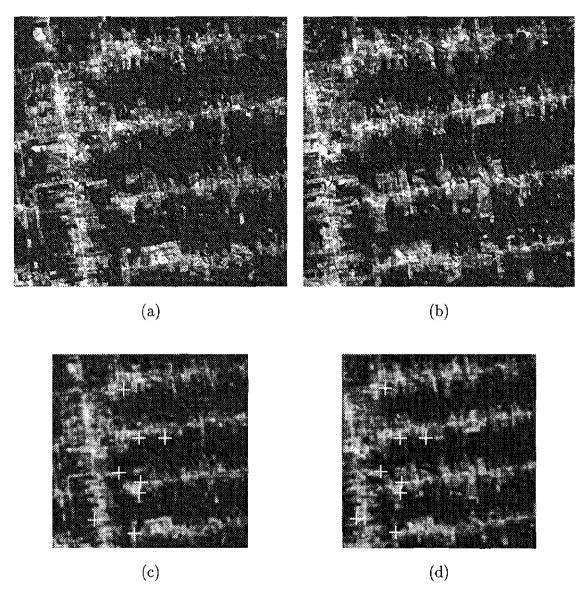
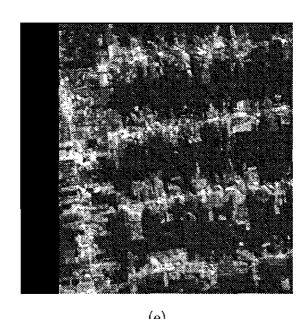


Fig. 7.12 - Imagens da Amazônia: (a) imagem referência - TM945AM; (b) imagem de ajuste - TM965AM; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM945AM e TM965AM, respectivamente; (e) Imagem TM965AM após a transformação final (L=2 e  $\beta=3$ ).



(e) Fig. 7.12 - Conclusão

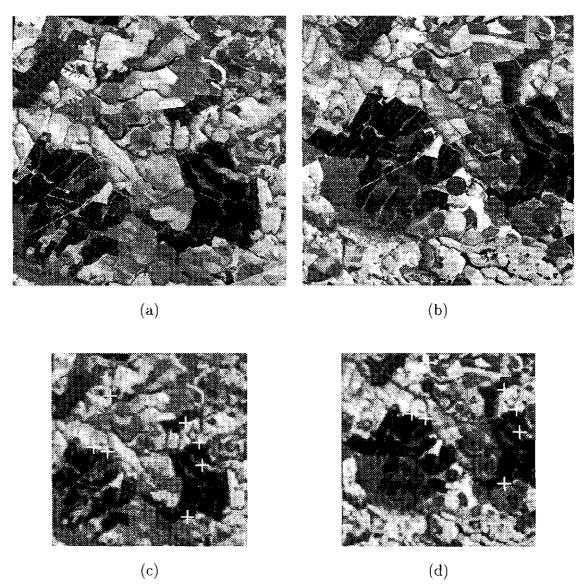


Fig. 7.13 - Imagens de área de agricultura: (a) imagem referência - TM945I; (b) imagem de ajuste - TM905I; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM945I e TM905I, respectivamente; (e) Imagem TM905I após a transformação final (L=2 e  $\beta=3$ ).

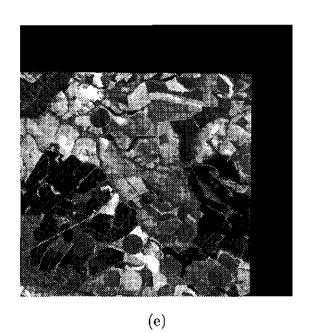


Fig. 7.13 - Conclusão

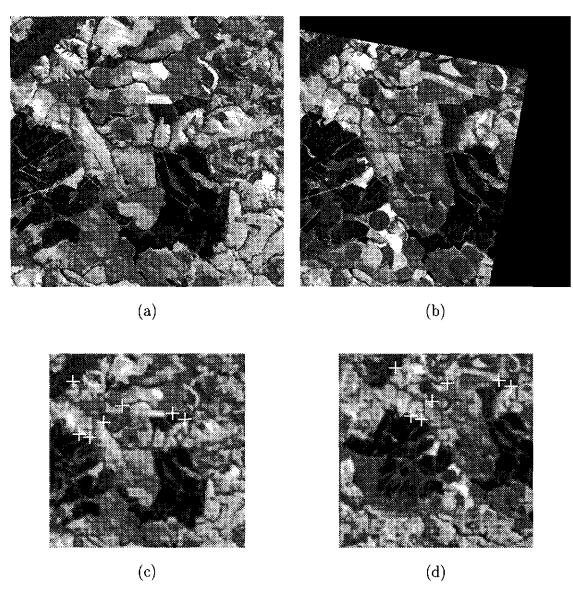


Fig. 7.14 - Imagens de área de agricultura: (a) imagem referência - TM945RI; (b) Imagem TM905I após a transformação final (L=2 e  $\beta=2$ ); (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM945RI e TM905I, respectivamente.

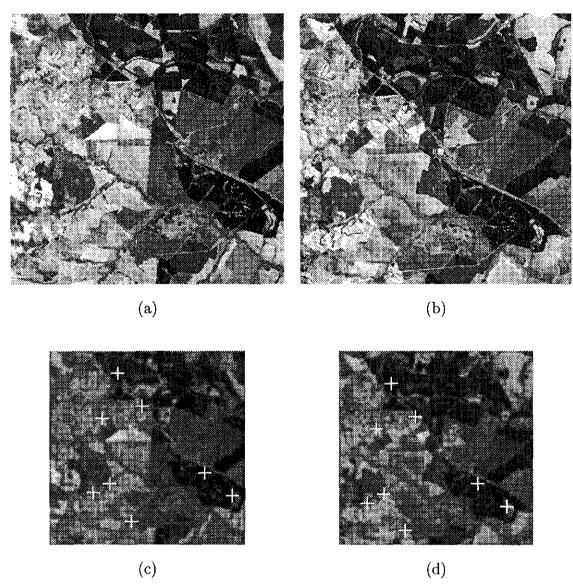
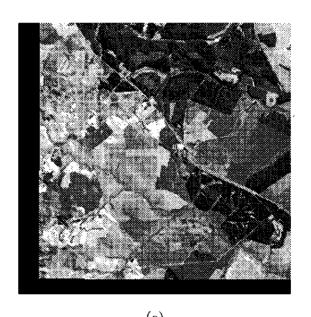


Fig. 7.15 - Imagens de área de agricultura(a) Imagem TM original tomada como referência (TM925AG);(b) Imagem TM original tomada como imagem de ajuste (TM945AG); (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM925AG e TM945AG, respectivamente; (e) Imagem TM945AG após a transformação final (L=2 e  $\beta=3$ ).



(e) Fig. 7.15 - Conclusão

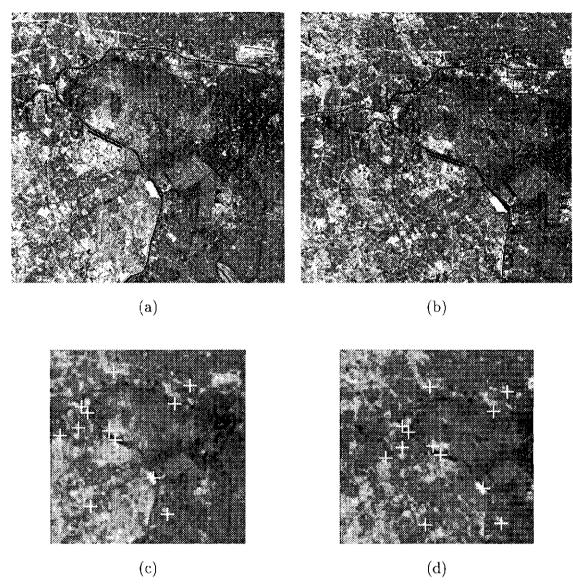
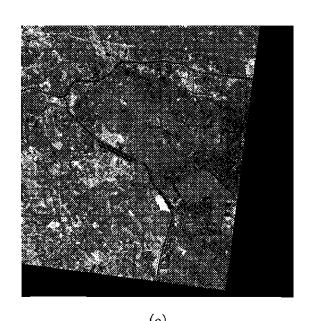


Fig. 7.16 - Imagens de área urbana:(a) Imagem TM original tomada como referência (TM944SP);(b) Imagem SPOT reamostrada tomada como imagem de ajuste (SP953SP); (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM944SP e SP953SP, respectivamente; (e) Imagem SP953SP após a transformação final.



(e) Fig. 7.16 - Conclusão

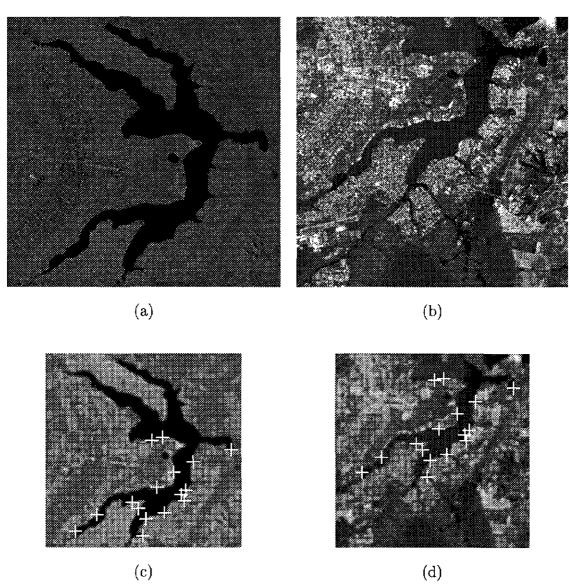
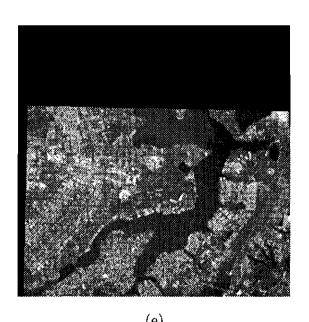


Fig. 7.17 - Imagens de área urbana:(a) Imagem TM original tomada como referência (TM944SP);(b) Imagem SPOT reamostrada tomada como imagem de ajuste (SP953SP); (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de TM944SP e SP953SP, respectivamente; (e) Imagem SP953SP após a transformação final.



(e) Fig. 7.17 - Conclusão

## 7.4 Experimento D

Para testar o potencial do algoritmo proposto no registro de imagens de radares, selecionamos dois pares de imagens multi-temporais do sistema JERS-1, JERS93/JERS96 e JERS93/JERS95 (vide Tabela 7.3). Antes do processo de registro as imagens foram processadas pelo filtro gama, com 10 iterações, para redução do ruído speckle (Ceasar, 1998). As imagens foram processadas usando a decomposição wavelet em L=2 níveis e os seguintes parâmetros:  $\beta=3$ ,  $w_c=13$ ,  $T_c=0.75$  e  $T_w=0.9$ . As imagens JERS foram registradas usando os modelos de similaridade e afim. Como nenhuma diferença foi observado nos resultados, mostramos apenas os resultados obtidos pelo o modelo de similaridade.

Os resultados do registro dos pares de imagens JERS93 e JERS96 e JERS93 e JERS95 filtradas são mostrados nas Figuras 7.18(e) e Figuras 7.19(e), respectivamente. A imagem JERS93 foi tomada como referência (Figuras 7.18(a) e 7.19(a)) nos dois testes e as imagens JERS96 e JERS95 nas Figuras 7.18(b) e 7.19(b), respectivamente, foram escolhidas como as imagens de ajuste. Os pontos de controle iniciais, sobrepostas às bandas LL das imagens originais no nível L, obtidos no registro dos pares de imagens JERS93 e JERS96 e JERS93 e JERS95 são mostrados, respectivamente, nas Figuras 7.18(c) e (d) e Figuras 7.19(c) e (d).

TABELA 7.3 - IMAGENS TESTES (RADAR)

Imagens	região	satélite	banda	data	área
JERS93	Amazônia	JERS-1	X	26/06/93	floresta
JERS96		JERS-1	X	13/08/96	
JERS93	Amazônia	JERS-1	X	26/06/93	floresta
JERS95		JERS-1	X	10/10/95	
JERS95B	Amazônia	JERS-1	X	10/10/95	floresta
JERS96B		JERS-1	X	13/08/96	

Finalmente, para mostrar o potencial do nosso algoritmo no registro de imagens maiores, apresentamos um resultado do registro de duas imagens multi-temporais JERS95B e JERS96B (vide Tabela 7.3), de tamanhos  $4000\times5920$  pixels, adquiridas pelo sistema de radar JERS-1. Da mesma forma anterior, antes do processo de registro as imagens foram filtradas pelo filtro gama com 10 iterações, para redução do ruído speckle. Após a filtragem, as imagens foram registradas usando os seguintes parâmetros:  $\beta=3$ , L=5,  $T_c=0$ , 75,  $w_c=13$  e  $T_w=0$ , 9. As Figuras 7.20(a) e (b) mostram a imagem de referência e a imagem de ajuste JERS96B, após o registro. As imagens nas Figuras 7.20(a) e (b) estão reamostradas. Por isso, não é possível ver com detalhes as diferenças provocadas pelas variações sazonais. Os pontos de controle sobrepostos às bandas LL das imagens de referência e de ajuste, no nível de resolução mais baixa, são mostrados nas Figuras 7.20(c) e (d), respectivamente.

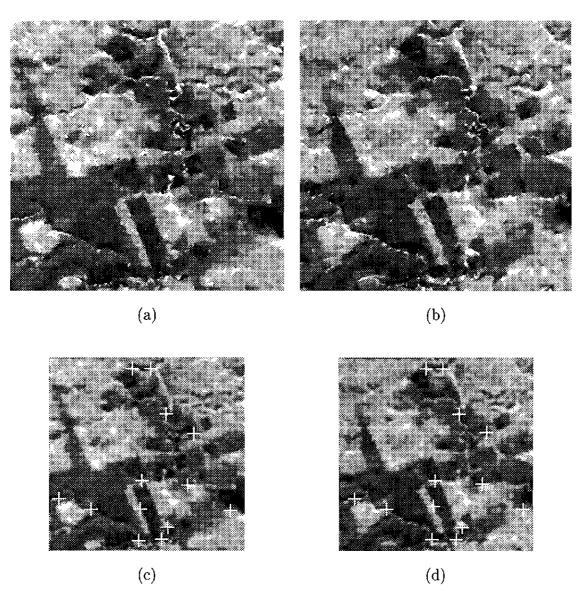
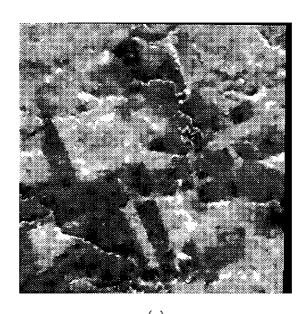


Fig. 7.18 - Registro das imagens JERS93 e JERS96; (a) Imagem de referência: JERS93; (b) Imagem de ajuste: JERS96; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de JERS93 e JERS96, respectivamente; (e) Imagem JERS96 após o registro.



(e) Fig. 7.18 - Conclusão

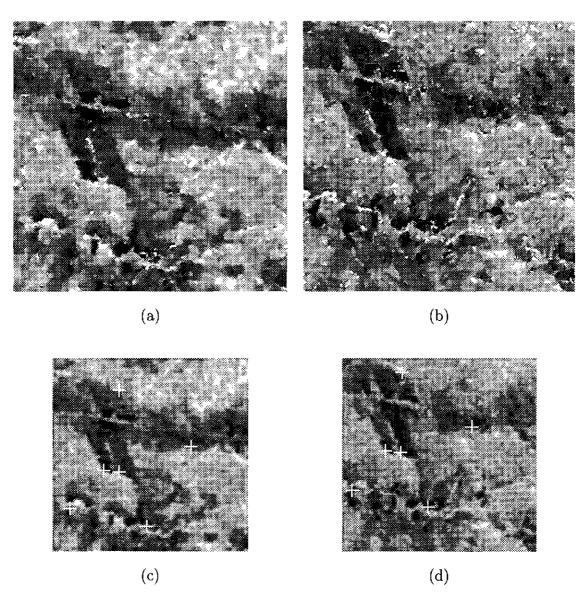


Fig. 7.19 - Registro das imagens JERS93 e JERS95:(a) Imagem de referência: JERS93;
(b) Imagem de ajuste: JERS95; (c) e (d) mostram os CPs iniciais sobrepostos
às bandas LL de JERS95 e JERS96, respectivamente; (e) Imagem JERS95
após o registro.

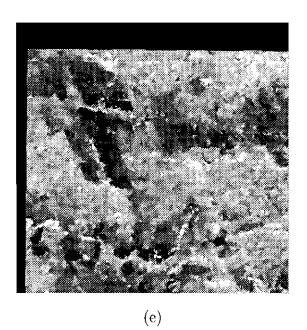


Fig. 7.19 - Conclusão

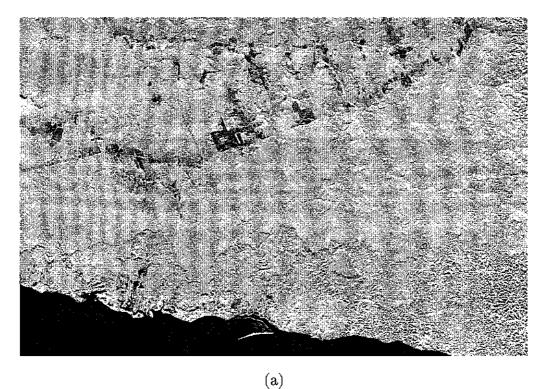
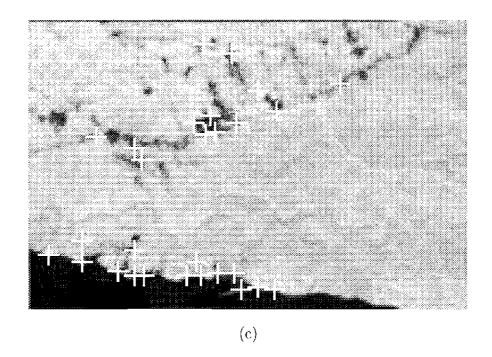


Fig. 7.20 - Registro das imagens JERS95B e JERS96B :(a) Imagem de referência: JERS95B; (b) Imagem de ajuste, JERS96B, após o registro; (c) e (d) mostram os CPs iniciais sobrepostos às bandas LL de JERS95B e JERS96B, respectivamente.



(d) Fig. 7.20 - Conclusão

## CAPÍTULO 8

## AVALIAÇÃO DOS RESULTADOS

Embora a maioria das imagens usadas nos experimentos tenha a mesma banda espectral ou bandas similares, as diferenças sazonais presentes nestas imagens tornam difícil o problema de registro automático. Antes do processo de registro, as imagens podem ser pré-processadas para melhorar o constraste ou para reduzir ruídos, conforme discutido no Capítulo 6. Neste experimento algumas imagens foram realçadas (transformação linear do histograma) para melhorar a qualidade visual ou filtradas para reduzir o ruído *speckle* das imagens de radar.

Para todas as imagens testes, processadas pelo algoritmo proposto, os resultados foram satisfatórios. Visualmente, comparamos os resultados sobrepondo a imagem de ajuste, registrada pelo algoritmo proposto, com a imagem de referência no sistema SPRING (Sistema de Processamento de Informações Georeferenciadas) (Câmara et al., 1996). Colocamos no canal R a imagem de ajuste corrigida e nos canais G e B a imagem de referência. Para uma análise visual dos resultados, as Figuras 7.21-7.26 mostram as composições coloridas das imagens registradas com as imagens de referência para seis pares de imagens testes. Podemos observar que feições tais como estradas e rios se sobrepõem satisfatoriamente, o que indica uma boa qualidade do registro.

Para medir quantitativamente o erro de registro usamos a raiz quadrada do erro médio quadrático (RMSE) definido na Equação 6.5. Afim de se obterem resultados independentes, foram selecionados um conjunto de pontos de controle diferentes daqueles usados no processo de registro automático das imagens, os quais chamaremos de pontos de teste. Estes pontos foram adquiridos manualmente no sistema ENVI (The Environment for Visualizing Images), que usa o sistema de coordenadas linha e coluna da imagem de forma similar ao nosso algoritmo (ENVI, 1997). Dada a função de transformação obtida através do registro automático e este conjunto de



Fig. 8.1 - Composição colorida das imagens TM945AM e TM925AM registrada.

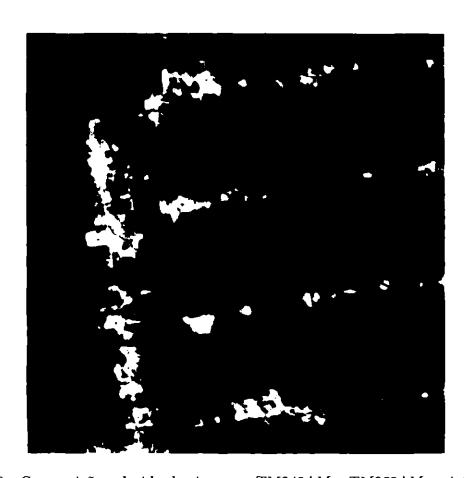


Fig. 8.2 - Composição colorida das imagens TM945AM e TM965AM registrada.



Fig. 8.3 - Composição colorida das imagens TM945I e TM905I registrada.



Fig. 8.4 - Composição colorida das imagens TM945RI e TM905I registrada.



Fig. 8.5 - Composição colorida das imagens TM925AG e TM945AG registrada.

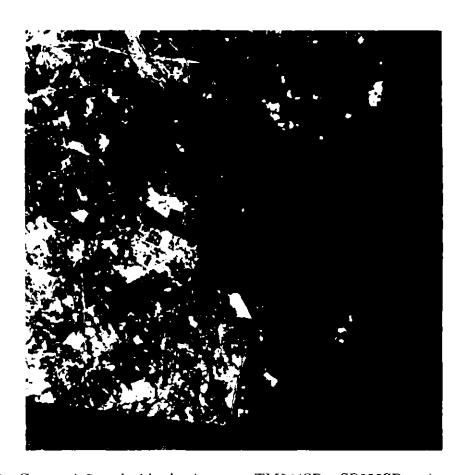


Fig. 8.6 - Composição colorida das imagens TM944SP e SP953SP registrada.

pares de pontos de controle (pontos de teste), o erro de registro RMSE (Equação 6.5) é calculado para cada par de imagens testes.

Para avaliar o desempenho e o tempo de processamento do algoritmo, as imagens foram registradas considerando dois casos: (1) processamento em múltiplas resoluções (L=2) e (2) processamento na resolução da imagem original. Por conveniência, referiremos ao processamento no nível de resolução da imagem original por L=0. A Tabela 8.1 mostra os valores dos parâmetros  $\beta$  e  $T_c$ , o número de pontos de controle obtidos no registro (#PCs), o erro RMSE (unidade de medida em pixels) e o tempo de processamento (medido em segundos) para os dois níveis de processamento. Os valores dos parâmetros  $w_c$  e  $T_w$  foram fixados, respectivamente, em 13 e 0,9 em todos os testes, com exceção do par de imagens JERS93/JERS95, para o qual  $w_c = 15$  em L=0.

Para os pares de imagens TM905I/TM945I e TM905I/TM945RI, o valor do parâmetro  $T_c$  foi fixado em 0,7 para levar em conta a baixa similaridade entre as imagens, provocada pelas variações sazonais. O parâmetro  $\beta$  foi também ajustado para gerar um maior número de pontos de controle e assim garantir o registro correto, considerando que o registro destas imagens é um pouco mais difícil.

O tempo de processamento foi computado em uma estação SUN-ULTRA 30, 300MHz e 640 MBytes de memória, em condições similares para todos os testes. Este tempo inclui todo o processo de registro, inclusive a leitura da imagem em disco, o cálculo da transformada wavelet das imagens e a interpolação bilinear para transformar a imagem de ajuste original. Os processos de realce de contraste e filtragem de ruído não estão incluídos.

Na Tabela 8.1 podemos observar que, para a maioria dos casos, o erro de registro para o caso L=2 foi menor do que 1 pixel, o que é um indicativo da boa precisão do registro. De um modo geral, os erros de registro foram semelhantes para os dois níveis de processamento, com exceção das imagens de radares. Como era esperado, o processamento em multi-resolução obteve melhor desempenho do que o proces-

**TABELA 8.1 -**ERROS DE REGISTRO (RMSE) E TEMPO DE PROCESSAMENTO

Imagens	Nível (L)	β	$T_c$	# PCs	RMSE (pixel)	Tempo (seg.)
TM925AM	0	3	0,75	56	0,754	1029
TM945AM	2	3	0,75	169	0,787	29
TM975AM	0	3	0,75	120	0,665	913
TM955AM	2	3	0,75	485	0,655	31
TM965AM	0	3	0,75	114	0,781	968
TM945AM	2	3	0,75	326	0,828	29
TM905I	0	3	0,7	85	1,158	817
TM945I	2	3	0,7	184	0,989	27
TM905I	0	2	0,7	52	0,861	2364
TM945RI	2	2	0,7	287	0,828	42
TM945AG	0	3	0,75	84	0,578	965
TM925AG	2	3	0,75	294	0,344	28
SP953SP	0	3	0,75	46	0,561	811
TM944SP	2	3	0,75	267	0,434	29
SP953DF	0	3	0,75	11	0,951	754
TM944DF	2	3	0,75	110	0,915	29
JERS93	0	3	0,75	16	1,582	569
JERS96	2	2	0,75	153	1,101	30
JERS93	0	3	0,75	5	2,349	1067
JERS95	2	2	0,75	79	0,669	26

samento no nível L=0 para o registro das imagens de radares, embora o erro de registro tenha sido maior do que 1 pixel para as imagens JERS93/JERS96. Por outro lado, a suavização da textura e ruído, nos níveis de resolução mais baixa, facilitou a obtenção de pontos de controle tornando possível o registro destas imagens com precisão de registro satisfatória.

Em relação ao tempo de processamento podemos observar, na última coluna da Tabela 8.1, que o registro em multi-resolução é mais eficiente que o processamento em L=0. Com precisões de registro semelhantes e em alguns casos até melhores, o registro em multi-resolução chega a ser dezenas de vezes mais rápido do que o processamento no nível 0. Note-se que esta diferença pode ser reduzida se alterarmos os valores do parâmetro  $\beta$ , cujos valores foram mantidos iguais para efeito de comparação.

Um outro experimento realizado foi a modificação do parâmetro  $\beta$  no caso de L=0, de tal forma que os tempos de processamento para os dois níveis, ficassem aproximadamente iguais. A Tabela 8.2 mostra os valores dos parâmetros do algoritmo usados,  $\beta$  e  $T_c$ , o número de pontos de controle obtidos, o erro RMSE e o tempo de processamento para o nível de processamento L=0. Nesta tabela, o símbolo \* indica que o registro não foi realizado com sucesso dada a restrição de tempo. Em dois casos (TM945AG/TM925AG e TM975AM/TM955AM), nos quais as imagens possuem área de sobreposição grande, os erros são pequenos e menor do que 1 pixel. No experimento TM975AM/TM955AM, a precisão de registro é praticamente igual ao do processamento em multi-resolução. Nos outros casos houve uma considerável perda da precisão do registro e um caso de registro incorreto devido à obtenção de pontos de controle incorretos, que resultaram em uma estimativa incorreta dos parâmetros da função de transformação.

Devido ao grande volume de dados do par de imagens JERS95B e JERS96B, não foi possível realizar o registro no nível L=0, sob as mesmas condições do que o nível L=2. O tempo de processamento ultrapassou cinco (5) dias e o programa foi abortado.

**TABELA 8.2 -**ERROS DE REGISTRO (RMSE) E TEMPO DE PROCESSAMENTO PARA O NÍVEL 0

Imagens	β	$T_c$	# PCs	RMSE (pixel)	Tempo (seg.)
TM925AM/TM945AM	3	0,75	5	1,068	22
TM975AM/TM955AM	6	0,75	11	0,665	24
TM965AM/TM945AM	6	0,75	3	1,43	22
TM905I/TM945I	5	0,7	12	1,382	52
TM905I/TM945RI	5	0,7	3	5,102	54
TM945AG/TM925AG	5	0,75	22	0,676	54
SP953SP/TM944SP	6	0,75	6	1,442	28
SP953DF/TM944DF	6	0,75	3	29,91	49
JERS93/JERS96	8	0,75	9	1,462	29
JERS93/JERS95	6	0,75	0	*	*

Para verificar o comportamento do erro de registro (RMSE) para diferentes níveis de processamento, o seguinte experimento foi realizado. Uma imagem Landsat-TM, da região amazônica, 3354x4096 pixels, banda 5, foi selecionada e transformada através da função de similaridade com os seguintes parâmetros: s = 0, 90;  $\theta = 10$ ;  $\Delta x = 200$  e  $\Delta y = -400$ . As imagens original e simulada foram registradas para 7 diferentes níveis de processamento (L=0,...,6).

O gráfico da Figura 8.7 mostra os valores dos erros de registro (RMSE) para estes níveis de processamento. A precisão de registro é melhor (erro menor) quanto maior é o número de níveis de processamento que as as imagens são registradas? não. Como podemos observar na Figura 8.7, os erros de registro nos níveis L=5 e L=6 são os maiores comparados aos outros níveis. O erro de registro final depende fortemente da estimativa inicial dos parâmetros da função de transformação, que por sua vez depende da boa precisão dos pontos de controle obtidos. Se a estimativa inicial dos parâmetros da função de transformação é muito grosseira, a precisão do registro final também é ruim. Daí vem a importância de se ter bons pontos de controle no registro inicial para se ter um registro final de boa precisão. Quando se processa as imagens em níveis de resolução muito baixa, a dificuldade de se obter um número de pontos de controle suficientes e de boa precisão é maior.

O gráfico da Figura 8.8 mostra o processo de refinamento do registro das imagens TM95 e TM97, ao longo dos diferentes níveis de resolução, partindo do nível 6 até o nível 0. No nível de resolução mais grosseira (6), quando se inicia o processo de registro, a função de transformação é estimada de maneira grosseira. À medida que o processamento caminha para os níveis de resolução progressivamente mais fina, a função de transformação vai se refinando. A Tabela 8.3 mostra os parâmetros da função e os erros de registro nos diferentes níveis de resolução para o nível de processamento L=6.

No caso de imagens de radar, as imagens JERS95B e JERS96B foram registradas usando a decomposição em 6 níveis e os seguintes parâmetros  $w_c = 13$ ,  $\beta = 1$ ,  $T_c = 0,75$  e  $T_w = 0,95$ . O tempo de processamento foi de aproximadamente 3 horas

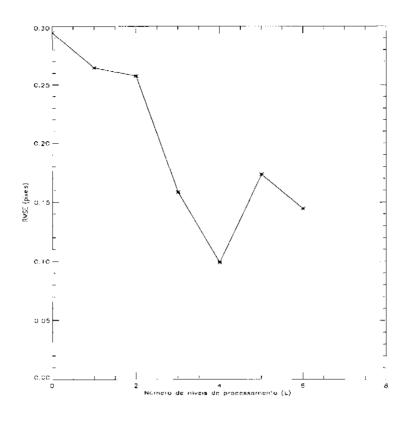


Fig. 8.7 - Erros de registro (RMSE) para diferentes níveis de processamento.

**TABELA 8.3** -PARÂMETROS DE TRANSFORMAÇÃO E ERROS (RMSE) PARA OS DIFERENTES NÍVEIS DE RESOLUÇÃO

L	$\theta$	s	$\Delta y$	$\Delta x$	RMSE
6	10,976438	0,865999	3,271483	-4,682739	81,207550
5	9,960503	0,895133	6,638279	-12,267748	8,380517
4	10,001248	0,898803	12,571225	-24,839426	2,665848
3	10,001767	0,899052	25,080223	-49,734436	2,286595
2	10,000772	0,899866	49,972927	-99,890846	0,511632
1	10,000525	0,899876	99,942017	-199,795151	0,494677
0	9,997562	0,900021	199,958817	-399,992249	0,144534

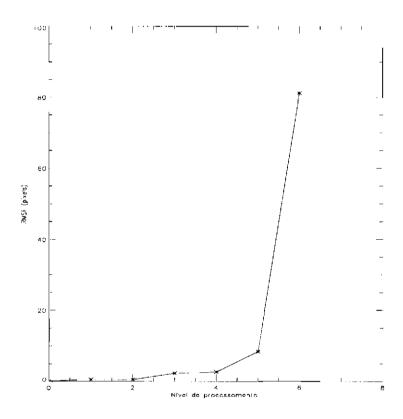


Fig. 8.8 - Erros de registro (RMSE) nos diferentes níveis de resolução (L=6), para as imagens Landsat-TM5.

com uma precião de registro de 1,103. O gráfico na Figura 8.9 mostra o processo de refinamento do registro ao longo dos diferentes níveis de resolução, partindo do nível de resolução mais baixa (6) até o nível 0. Pode-se observar que na passagem do nível 1 para o nível 0, a precisão do registro piorou, ao invés de melhorar. Este problema foi, provavelmente, provocado pelo o efeito do ruído speckle e textura das imagens. Além disso, as imagens apresentam muitos padrões que se repetem em toda a extensão da imagem, que dificultam o casamento correto das feições. Uma das soluções encontradas para este problema foi o uso, no último nível de processamento, de uma janela de correlação de tamanho maior do que a definida no início do registro. Foram testados os tamanhos de janela 21 e 41, e a janela 41x41 obteve o melhor resultado. Estes valores foram escolhidos nos testes porque alguns outros experimentos, com estas imagens, já haviam sido realizados anteriormente. Com esta modificação no algoritmo, o erro de registro passou a ser de 0,886, menor do que o do resultado anterior mas ainda um pouco maior do que o do nível 1. A Tabela 8.4 mostra os valores dos erros de registro para cada nível de resolução após a modificação no algoritmo, para o nível de processamento L=6.

**TABELA 8.4** -ERROS (RMSE) PARA OS DIFERENTES NÍVEIS DE RESOLUÇÃO (L=6)

$\lfloor L \rfloor$	RMSE
6	12,649131
5	8,284471
4	3,713292
3	2,214443
2	0,895715
1	0,829640
0	0,886410

Finalmente, para efeito de comparação, testamos o método baseado em feições pro-

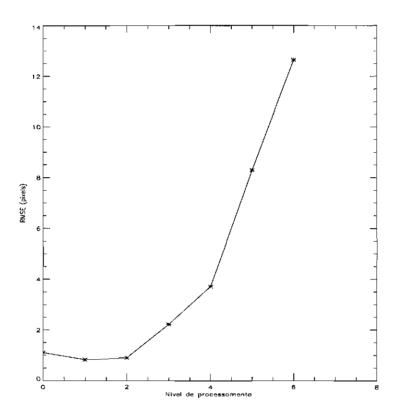


Fig. 8.9 - Erros de registro (RMSE) nos diferentes níveis de resolução (L=6), para as imagens JERS.

posto por Li et al (1995). O método de Hui Li falhou na maioria dos casos, porque não existe uma consistência dos contornos dos objetos nas imagens multi-temporais a serem registradas. Nos casos em que o método teve sucesso, poucos contornos fechados foram casados (em torno de 3) e os valores do erro de registro foram altos: 2,6 (TM905I/TM945RI), 1,9 (TM945AG/TM925AG) e 4,5 (SP953SP/TM944SP).

### CAPÍTULO 9

# CONCLUSÕES E DIREÇÕES PARA PESQUISAS FUTURAS

### 9.1 Resumo e Conclusões

Nesta tese apresentamos um método para registrar automaticamente imagens de sensoriamento remoto. O método usa a estratégia de multi-resolução baseada na transformada wavelet, o que torna o algoritmo eficiente, rápido e de fácil implementação. A decomposição wavelet gera sub-imagens com conteúdo de informações diferentes, as quais são usadas para extrair feições significativas nas imagens (bandas de altas freqüências) e também para realizar o casamento entre elas (bandas de baixas freqüências). A estrutura piramidal é usada para realizar o casamento nos diferentes níveis de resolução, usando-se o resultado obtido nos níveis de resolução mais baixa para refinar os parâmetros nos níveis de resolução progressivamente mais alta.

Para se obter as feições usadas no processo de casamento, propomos um método automático de extração de feições baseado nos máximos locais do módulo da transformada wavelet. O procedimento, basicamente com um parâmetro, definido pelo usuário, gera automaticamente um conjunto de feições consistentes e significativas, que podem ser usadas como pontos de controle no processo de registro.

A geração de pontos de controle corretos no processo de casamento inicial é a fase crucial de todo o algoritmo. Uma estimativa errada dos parâmetros do modelo de distorção inviabiliza o registro correto. Para selecionar um conjunto de pontos de controle corretos, propomos dois métodos de verificação de consistência dos casamentos. O primeiro método é baseado na hipótese de que as distâncias entre dois pontos nas imagens são preservadas e, portanto, é mais adequado para imagens cuja geometria interna é modelada por uma distorção rígida. O segundo, usa uma busca combinatória, entre os N casamentos obtidos na fase inicial, para selecionar os casamentos corretos. Embora este exija um esforço computacional maior do que o

primeiro método, ele possui a vantagem de ser independente do modelo de distorção das imagens.

O algoritmo proposto neste trabalho resolve, com algumas restrições, o grande problema que é a escolha dos pontos de controle, que demanda pessoal especializado, tempo e que, sobretudo, é uma tarefa desgastante. O algoritmo é fácil de ser aplicado porque ele precisa basicamente de dois parâmetros: um que define o número de níveis da decomposição wavelet e outro que define o número de feições extraídas usadas para gerar os pontos de controle. Em termos de complexidade computacional o algoritmo é eficiente. O tempo de processamento depende da imagem, mas de um modo geral gasta menos que um minuto em uma estação SUN-ULTRA 30 para um par de imagens de 512x512 pixels. Comparado ao processamento realizado no nível de resolução 0, o processamento em múltiplas resoluções chega a ser dezenas de vezes mais rápido, com a mesma ou melhor precisão de registro.

Os casos de registro de imagens mais difíceis, mencionados na Seção 2.4, a saber: registro de imagens com resoluções diferentes, registro de imagens multi-temporais e de imagens afetadas por ruídos são solucionados de alguma forma neste trabalho. As imagens com resoluções diferentes são reamostradas para um mesmo (similar) tamanho de pixel antes do processo de registro. No caso de registro de imagens multi-temporais, o uso de feições pontuais no processo de casamento, ao invés de contornos dos objetos, torna o registro mais robusto considerando que feições pontuais correspondentes nestas imagens são mais fáceis de serem encontradas. O processo de filtragem e a estratégia de multi-resolução reduzem significativamente o problema do ruido speckle nas imagens de radares, tornando factível o registro automático destas imagens.

A principal restrição do algoritmo é que ele é mais apropriado para registrar imagens com respostas espectrais similares, pois a informação espectral é utilizada no processo de casamento de feições. O registro de imagens obtidas por sensores diferentes pode ser realizado com sucesso desde que as bandas espectrais sejam escolhidas adequadamente, como foi mostrado neste trabalho, para o caso do registro de imagens

#### TM-4 e SPOT-3.

Vários resultados experimentais e medidas de desempenho foram apresentados para ilustrar o desempenho do algoritmo de registro proposto. Os resultados foram avaliados através do erro de registro, medido pelo critério da raiz quadrada do erro médio quadrático e através de uma análise visual pela sobreposição das imagens registradas. Os resultados foram bastante satisfatórios.

Portanto, os resultados experimentais apresentados validam a metodologia proposta neste trabalho.

#### 9.2 Futuras Pesquisas

Para pesquisas futuras, sugerimos os seguintes tópicos:

- a formalização da escolha dos parâmetros do algoritmo de registro;
- a otimização da escolha das wavelets para o problema de registro de imagens;
- o estudo de critérios ou medidas de avaliação da precisão do registro;
- Estudo de metodologias que garantam uma boa distribuição dos pontos de controle obtidos pelo método proposto.
- Integração do algoritmo ao sistema SPRING;

Uma sugestão para a escolha dos parâmetros, seria a inclusão no algoritmo de um módulo inteligente. Este módulo seria capaz de examinar algumas propriedades das imagens fornecidas pelo usuário tais como por exemplo, o grau de sobreposição e similaridade entre as imagens, intervalo de tempo de aquisição entre as imagens, ou informações contidas nos dados e previamente conhecidas (por exemplo, tamanho das imagens e rotação). Estas informações seriam usadas para estimar os valores dos parâmetros. O tamanho da janela de correlação poderia ser determinada através do procedimento usado na Seção 5.3.1, onde calcula-se a taxa de acertos (casamentos corretos) em função do tamanho da janela de correlação.

A escolha das wavelets tem sido um tópico muito estudado em todas as áreas de aplicação que usam as wavelets. Para o problema de registro têm sido estudados o uso de wavelets invariantes à translação (TIW) e wavelets compactas e a influência do uso de diferentes wavelets na qualidade do registro (Chettri et al.,1997; Deubler e Oliver, 1997). Para o algoritmo de registro proposto nesta tese, seria de interesse o estudo da influência do uso de diferentes wavelets na precisão do registro.

O registro de imagens de sensoriamento remoto é um processo de aprendizagem iterativo onde a escolha dos pontos de controle e a forma de avaliação da precisão de registro devem ser realizados de forma criteriosa e com metodologias adequadas. Assim, acreditamos que o estudo de critérios ou medidas de avaliação da precisão do registro é uma área de pesquisa que deve ser mais investigada. Buiten e Putten (1997) propoēm uma metodologia para analisar o comportamento geométrico dos "resíduos" a partir da hipótese da sua aleatoridade, usando alguns testes estatísticos. Estes testes poderiam ser realizados para a avaliação do registro realizado pelo nosso algoritmo.

Uma boa distribuição dos pontos de controle é um dos requisitos para se obter um registro de boa qualidade. Esta é uma linha de pesquisa que ainda tem muito o que ser feito.

Dada a importância tecnológica deste trabalho, haja vista a não disponibilidade de técnicas de registro automático nos atuais sistemas de processamento de imagens, espera-se integrar este algoritmo ao sistema SPRING disponibilizando-o à comunidade de sensoriamento remoto.

# REFERÊNCIAS BIBLIOGRÁFICAS

- Abbasi-Dezfouli, M.; Freeman, T. G. Patch matching in stereo-images based on shape. In: The International Society for Optical Engineering. Spatial information from digital photogrammetry and computer vision. Bellingham, Wash.: SPIE, 1994. p. 1-8. (SPIE Proceedings, v. 2357)
- Alves, D. S.; Meira Filho, L. G.; d'Alge, J. C. L.; Mello, E. K.; Moreira, J. C.; Medeiros, J. S. The Amazonia information system. In: International Congress for Photogrammetry and Remote Sensing, 17., Washington, DC, 1992. International Archives of Photogrammetry and Remote Sensing. Bethesda, MD: ISPRS, 1992. V. 29, Part B6, p. 259-266.
- Antonini, M.; Berlaud, M.; Mathieu, P.; Daubechies, I. Image coding using wavelet transform. **IEEE Transactions on Image Processing**, v. 1, n. 2, p. 205-220, Apr. 1992.
- Anuta, P. E. Spatial registration of multispectral and Multitemporal Digital Imagery using Fast Fourier Transform Techniques. IEEE Transactions on Geoscience Electronics, v. 8, n. 4, p. 353-368, Oct. 1970.
- Ballard, D. H.; Brown, C. M. Computer and vision. New Jersey: Prentice-Hall, 1982.
- Banon, G. J. F.; Barrera, J. Minimal representations for translation-invariant set mappings by mathematical morphology. **SIAM Journal Applied Mathematic**, v. 51, n. 6, p. 1782-1798, Dec. 1991.
- Banon, G. J. F.; Barrera, J. Decomposition of mappings between complete lattices by mathematical morphology. **Signal Processing**, v. 30, n. 3, p. 299-327, Feb. 1993.

- Banon, G. J. F. Characterization of translation-invariant elementary morphological operators between gray-level images. São José dos Campos: INPE, 1995. (INPE-5616-RPQ/671).
- Banon, G. J. F.; Faria, S. D. Morphological approach for template matching.
  In: Brazilian Symposium on Graphic Computation and Image Processing, 10.,
  Campos de Jordão, 1997. Proceedings. Los Alamitos: IEEE Computer Socitey. p. 171-178.
- Banon, G. J. F. Formal introduction to digital image processing. [online]. <a href="http://hermes.dpi.inpe.br:1905/rep/dpi.inpe.br/banon/1998/07.02.12.54">http://hermes.dpi.inpe.br:1905/rep/dpi.inpe.br/banon/1998/07.02.12.54</a>. Apr. 1998b.
- Banon, G. J. F.; Barrera, J. Bases da Morfologia Matemática para a Análise de Imagens Binárias. [online]. <a href="http://hermes.dpi.inpe.br/1905/">http://hermes.dpi.inpe.br/1905/</a> rep/dpi.inpe.br/banon/1998/06.30.17.56>. Apr. 1998a
- Barnea, D. I.; Silverman, H. F. A class of algorithms for fast digital image registration. IEEE Transactions on Computers, v. 21, n. 2, p. 179-186, Feb. 1972.
- Bartl, R.; Petrou, M.; Christmas, W. J.; Palmer, P. L. On the automatic registration of cadastral maps and Landsat TM images. In: The International Society for Optical Engineering. Image and signal processing for remote sensing III. Bellingham, Wash.: SPIE, 1996. p. 9-20. (SPIE Proceedings, v. 2955)
- Batista, G. T.; Medeiros, J. S.; Mello, E. M. K.; Moreira, J. C.; Bins, L. S. A new approach for deforestation assessment. In: International Symposium on Resource and Environmental Monitoring. Rio de janeiro, 1994. Proceedings.
  São José dos Campos: INPE, 1994. v. 30, p. 170-174.
- Bijaoui, A.; Giudicelli, M. Optimal Image Addition Using the Wavelet Transform. Experimental Astronomy, v. 1, n. 6, p. 347-363, 1991.

- Bins, L. S. Uso de casamento estrutural para registro de imagens de satélite. São José dos Campos. 106p. (INPE-4734-TDL/345). Dissertação (Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, 1988.
- Bins, L. S.; Fonseca, L. M. G.; Erthal, G. J.; Ii, F. A. M. Satellite Imagery segmentation: region growing. [CD-ROM]. In: Simpósio Brasileiro de Sensoriamento Remoto. 8., Salvador, 1996. Anais. São Paulo: Image Multimidia, 1996. Seção de Comunicações Técnico-Científicas.
- Brown, L. G. A survey on image registration techniques. **ACM Computing Surveys**, v. 24, n. 4, p. 325-376, Dec. 1992.
- Brum, E. R. Integração dos canais multiespectrais e pancromáticos do sensor HRV SPOT para obtenção de composições coloridas com resolução espacial próxima a 10 m. São José dos Campos. 113p. (INPE-5035-TDL/403). Dissertação (Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, 1989.
- Brunelli, R.; Messelodi, S. Robust Estimation of Correlation with Applications to Computer Vision. **Pattern Recognition**, v. 28, n. 6, p. 833-841, Jun. 1995.
- Buiten, H. J.; Putten, B. Quality assessment of remote sensing image registrationanalysis and testing of control points residuals. **ISPRS Journal of Pho**togrammetry and Remote Sensing, v. 52, n. 2, p. 57-73, Apr. 1997.
- Burrough, P. A. Principles of geographical information systems for land Resources assessment. New York: Oxford University Press, 1987.
- Burt, P. J.; Adelson, E. H. The Laplacian pyramid as a compact image code.

  IEEE Transactions on Communications, v. COM-31, n. 4, p. 532-540,

  Apr. 1983.

- Busse, L. J.; Crimmins, T. R.; Fienup, J. R. A model based approach to improve the performance of the geometric filtering speckle reduction algorithm. In: IEEE Ultrasonics Symposium, Scattle, Wash., 1995. **Proceedings**. New York: IEEE Computer Society, 1995. v. 2, p. 1353-1356.
- Câmara, G. Registro automático de imagens de satélite. In: Simpósio Brasileiro de Sensoriamento Remoto, 4., Gramado, 1986. Anais. São José dos Campos: INPE, 1986. v. 1, p. 669-671.
- Câmara, G.; Mascarenhas, N. D. A. Methods for Image Interpolation through FIR Filter Design Techniques. In: International Conference on Acoustics, Speech and Signal Processing, Boston, Mass., 1983. **Proceedings**. p. 391-394.
- Câmara, G.; Ii, F. A. M. Registro automático de imagens de satélites ambientais. São José dos Campos: INPE, 1985. (INPE-3578-PRE/781).
- Câmara, G.; Souza, R. C. M.; Freitas, U. M; Garrido, J. SPRING: Integrating Remote Sensing and GIS by Object Oriented data Modeling. Computers and Graphics, v. 20, n. 3, p. 395-403, May-June 1996.
- Candeias, A. L. B. Uso da teoria de estimação bayesiana na fusão de dados de satélites. São José dos Campos. 133 p. (INPE-5457-TDI/499). Dissertação (Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, 1992.
- Canny, J. F. A computational approach to edge detection. IEEE Transactions on Pattern Anal. and Machine Intell., v. 8, n. 6, p. 679-698, Nov. 1986.
- Carlson, A. B. Sistemas de comunicação. New York: McGraw-Hill, 1981.
- Carper, T. W.; Lillesand, T. M.; Kiefer, R. W. The use of intensity-huc-saturation transformation for merging panchromatic and multispectral image Data. Photogrammetric engineering and Remote Sensing, v. 56, n. 4, p. 459-467, Apr. 1990.

- Caves, G. R.; Harley, P. J.; Quegan, S. Registering SAR images to digital map data using a template matching technique. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1991. **Proceedings**. v. 4, p. 1429-1431.
- Chandra, D. V. S. Fast registration of multispectral images using orthogonal transform features. In: Midwest Symposium on Circuits and systems, 35., Washington, D.C., 1992. **Proceedings**. v. 1, p. 703-706.
- Chang, T.; Jay Kuo, C. C. Texture analysis and classification with tree-structured wavelet transform. IEEE Transactions on Image Processing, v. 2, n. 4, p. 429-441, Oct. 1993.
- Chapman, Y. C. S.; Mahani, R. B. A new technique to extract range information from stereo images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 7, p. 768-773, July 1989.
- Chen, C. H.; Lee, J. S.; Sun, Y. N. Wavelet transformation for gray-level corner detection. **Pattern Recognition**, v. 28, n. 6, p. 853-861, Jun. 1995.
- Chen, M.; Kanade, T.; Pomerleau, D.; Rowley, H. A. Anomaly Detection through Registration.
  In: Image Registration Workshop, 1., Greenbelt, MA, 1997.
  Proceedings.
  Greenbelt: NASA Goddard Space Flight Center, 1998. p. 201-210.
- Cheong, C.; Aizawa, K.; Saito, T.; Hatori, M. Subband image coding with biorthonormal wavelets. **EICE Transactions Fundamentals of Eletronics, Communications and Computer Sciences**, v. E75-A, p. 871-881, July 1992.
- Chettri, S.; Moigne, J.; Campbell, W. Scope and applications of translation invariant wavelets to image registration. In: Image Registration Workshop, 1.,
  Greenbelt, MA, 1997. Proceedings. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 29-33.

- Chui, C. An Introduction to Wavelets. New York: Academic Press, 1992.
- Cideciyan, A. V.; Jacobson, S. G.; Kemp, C. M.; Knighton, R. W.; Nagel, J. H. Registration of high resolution images of the retina. In: The International Society for Optical Engineering. Medical imaging VI: image processing. Bellingham, Wash.: SPIE, 1992. p. 310-322. (SPIE Proceedings, v.1652)
- Corvi, M.; Nicchiotti, G. Multiresolution Image Registration. In: International Conference on Image Processing, Los Alamitos, CA, 1995.
   Proceedings.
   Los Alamitos: IEEE Computer Society Press, 1995. v. 3, p. 224-227.
- Crimmins, T. R. Geometric filter for speckle reduction. Applied Optics, v. 24, n. 10, p. 1438-1443, May 1985.
- d'Alge, J. C. L. Correção Geométrica. (Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1998). Comunicação pessoal.
- Dana, K.; Anandan, P. Registering of visible and infrared images. In: International Society for Optical Engineering. Architecture, Hardware and forward-looking infrared issues in automatic target recognition. p. 1-11. Bellingham, Wash.: SPIE, 1993. (SPIE Proceedings, v. 1957)
- Dare, P. M.; Dowman, I. J. An automated procedure for registering SAR and optical imagery based on feature matching. In: The International Society for Optical Engineering, Microwave Sensing and Synthetic Aperture Radar. Bellingham, Wash.: SPIE, 1996. p. 140-151. (SPIE Proceedings, v. 2958)
- Dare, P. M.; Ruskoné, R.; Dowman, I. J. Algorithm Development for the Automatic Registration of Satellite Images. In: Image Registration Workshop, 1.,
  Greenbelt, MA, 1997. Proceedings. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 83-88.
- Decastro, E.; Morandi, C. Registration of translated and rotated images using finite Fourier-Transforms. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 9, n. 5, p. 700-703, Sep. 1987.

- Deubler, J.; Olivo, J. C. A wavelet-based multiresolution method to automatically register images. **Journal of Mathematical Imaging and Vision**, v. 7, n. 3, p. 199-209, Jun. 1997.
- Djamdji, J. P.; Bijaoui, A. Disparity Analysis: A wavelet transform approach.

  IEEE Transactions on Geoscience and Remote Sensing, v. 33, n. 1, p. 67-76, Jan. 1995.
- Djamdji, J. P.; Bijaoui, A.; Maniere, R. Geometrical registration of images: the multiresolution approach. **Photogrammetric Engineering and Remote Sensing**, v. 59, n. 5, p. 645-653, May 1993.
- Donze, M.; Bijlsma, R. J.; Buiteveld, H.; Clevers, J. G. P. W.; Kasteren, H. W. J.;
  Meulstee, C.; Pellemans, A. H. J. M.; Looyen, W. J.; Swol, R. W. CAESAR:
  performance and first evaluation of application possibilities.
  Delft:
  Netherlands Remote Sensing Board (BCRS), 1989. (Report BCRS-89-06).
- Duda, R. O.; Hart, P. E. **Pattern classification and scene analysis**. New York: John Wiley & Sons, 1973.
- Environment for Visualizing Images (ENVI). **Users'guide**. Boulder, 1997. v. 1-2.
- Environment for Visualizing Images (ENVI). Gamma filter: Comparison with standard filters. [on line]. <a href="http://www.nasoftware.co.uk/caesar/compare.html">http://www.nasoftware.co.uk/caesar/compare.html</a>. Mar. 1999.
- Erthal, G. J. Registro translacional de imagens: métodos correlacionais e sequenciais. São José dos Campos. 120 p. (INPE-3456-TDL/189). Dissertação (Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, 1985.
- Erthal, G.J.; Velasco, F.R.D.; Mascarenhas, N.D.A. Um Sistema para registro translacional de segmentos de imagens Landsat. São José dos Campos: INPE, 1983. (INPE-2785-PRE).

- Fliege, N. J. Multirate Digital Signal Processing: multirate systems, filter banks, wavelets. New York: John Wiley & Sons, 1994.
- Flusser, J. An adaptive method for image registration. **Pattern Recognition**, v. 25, n. 1, p. 45-54,1992.
- Flusser, J.; Suk, T. A Moment-based Approach to registration of images with affine geometric distortion. IEEE Transactions on Geoscience and Remote Sensing, v. 32, n. 2, p. 382-387, Mar. 1994.
- Fonseca, L. M. G. Restauração e interpolação de imagens do satélite Landsat. São José dos Campos. 148 p. (INPE-6628-TAE/30). Dissertação (Mestrado em Engenharia Eletrônica) Instituto Tecnológico da Aeronáutica, 1988.
- Fonseca, L. M. G.; Manjunath, B. S. Registration Techniques for Multisensor Remotely Sensed Imagery. Photogrammetric Engineering and Remote Sensing, v. 62, n. 9, p. 1049-1056, Sept. 1996.
- Fonseca, L. M. G.; Costa, M. H. Automatic registration of satellite images. In: Brazilian Symposium on Graphic Computation and Image Processing, 10., Campos de Jordão, 1997. Proceedings. Los Alamitos: IEEE Computer Society, 1997. p. 219-226.
- Fonseca, L. M. G.; Costa, M. H.; Manjunath, B. S. and Kenny, C. Automatic registration of satellite imagery. In: Image Registration Workshop, 1., Greenbelt, MA, 1997.
  Proceedings. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 13-27.
- Fonseca, L. M. G.; Prasad, G. S. S. D.; Mascarenhas, N. D. A. Combined interpolation-restoration of Landsat images through FIR filter design techniques. International Journal Remote Sensing, v. 14, n. 3, p. 2547-2561, Sept. 1993.

- Foresti, C.; Oliveira, M. L. N.; Niero, M.; Parreira, E. M. M. F. O uso da técnica de registro de imagens para o monitoramento do crescimento urbano. São José dos Campos: INPE, 1980. (INPE-8108-PRE/108).
- Frery, A. C.; Sant'Anna, S. J. S.; Mascarenhas, N. D. A.; Bustos, O. H. Robust inference techniques for speckle noise reduction in 1-look amplitude SAR images.

  Applied Signal Processing, v. 4, n. 2, p. 61-76, 1997.
- Gonzalez, R. C.; Woods, R. E. **Digital image processing**. New York: Addisonwesley, 1992.
- Goshtasby, A. Piecewise linear mapping functions for image registration. Pattern Recognition, v. 19, n. 6, p. 459-466, 1986.
- Goshtasby, A. Piecewise cubic mapping functions for image registration. **Pattern recognition**, v. 20, n. 5, p. 525-533, 1987.
- Goshtasby, A. Registration of Images with geometric distortions. **IEEE Transactions on Geoscience and Remote Sensing**, v. 26, n. 1, p. 60-64, Jan. 1988a.
- Goshtasby, A. Image registration by local approximation methods. Image and Vision Computing, v. 6, n. 4, p. 255-261, Nov. 1988b.
- Goshtasby, A.; Stockman, G. C. Point pattern matching using convex hull edges. IEEE Transactions on Systems Man and Cybernetics, v. 15, n. 5, p. 631-637, Sep. 1985.
- Goshtasby, A.; Stockman, G. C.; Page, C. V. A region based approach to digital image registration with subpixel accuracy. IEEE Transactions on Geoscience and Remote Sensing, v. 24, n. 3, p. 390-399, May 1986.
- Granlud, G. H. Fourier preprocessing for hand print character recognition. IEEE Transactions on Computers, v. C-21, n. 2, p. 195-201, Feb. 1972.

- Graps, A. An introduction to wavelets. **IEEE Computational Science and Engineering**, v. 2, n. 2, p. 50-61, Summer 1995.
- Guindon, B. Automated Control Point Acquisition in radar-optical image registration. Canadian Journal of Remote Sensing, v. 11, n. 1, p. 103-112, July 1985.
- Heijmans, H. J. A. M. Morphological image operators. Boston: Academic Press, 1994.
- Hewer, G.; Kenney, C.; Manjunath, B. S.; Schyoen, R. Peer Group Image Processing. Santa Barbara: University of California, Santa Barbara (UCSB), 1997. (ECE Technical Report TR # 97-02).
- Jaillon, P.; Montanvert, A. Image mosaicking applied to three-dimensional surfaces.
   In: IAPR International Conference on Pattern Recognition, 12.,
   Jerusalem, 1994.
   Proceedings.
   Los Alamitos: IEEE Comput. Soc.
   Press, 1994.
   v. 1, p. 253-257.
- Jensen, A. S.; Lindvold, L.; Rasmussen, E. Transformation of image positions, rotations, and sizes into shift parameters. Applied Optics, v. 26, n. 9, p. 1775-1781, May 1987.
- Kass, M.; Witkin, A.; Terzopoulos, D. Snakes active contour models. In: International Conference on Computer Vision, Washington, DC, 1., 1987. Proceedings. Washington: Computer Society Press of the IEEE, 1987. p. 259-268.
- Kher, A.; Mitra, S. Registration of noisy SAR imagery using morphological feature extractor and 2-D cepstrum. In: The International Society for Optical Engineering. Applications of digital image processing XV. Bellingham, Wash.: SPIE, 1993.p. 281-291. (SPIE Proceedings, v. 1771)
- Khosravi, M.; Schafer, R. W. Template Matching Based on a Grayscale Hit-or-Miss Transform. IEEE Transactions on Image Processing, v. 5, n. 6, p. 1060-1066, Jun. 1996.

- Kitchen, L.; Rosenfeld, A. Gray-level corner detection. Pattern Recognition Letters, v. 1, n. 2, p. 95-102, Dec. 1982.
- Kuglin, C. D.; Hines, D. C. The phase correlation image alignment method. In: International Conference on Cybernetics and Society IEEE, New York, 1975.
  Proceedings. New York: Institute of Electrical and Electronics Engineers, 1975. p. 163-165.
- Laine, A.; Fan, J. An adaptative approach for texture segmentation by multi-channel wavelet frames. In: The International Society for Optical Engineering.
  Mathematical imaging: wavelet applications in signal and image processing. Bellingham, Wash.: SPIE, 1993. p. 288-299. (SPIE Proceedings, v. 2034)
- Lee, J. S.; Sun, Y. N.; Chen, C. H.; Tsai, C. T. Wavelet based corner detection.

  Pattern Recognition, v. 26, n. 6, p. 853-865, Jun. 1993.
- Li, H.; Zhou, Y. Automatic visual/IR image registration. Optical Engineering, v. 35, n. 2, p. 391-400, Feb. 1996.
- Li, H.; Zhou, Y. A wavelet based edge detector. IEEE Transactions on Image Processing, v. 4, n. 3, p. 161-164, Mar. 1995a.
- Li, H.; Manjunath, B. S.; Mitra, S. K. A contour-based approach to multisensor image registration. IEEE Transactions on Image Processing, v. 4, n. 3, p. 320-334, Mar. 1995.
- Loew, M. H. Issues in Multimodality Medical Image Registration. In: Image Registration Workshop, 1., Greenbelt, MA, 1997. **Proceedings**. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 1-7.
- Maes, F.; Vandermeulen, D.; Mar.al, G.; Suetens, P. Fast Multimodality Image Registration Using Multiresolution Gradient-based Maximization of Mutual Information. In: Image Registration Workshop, 1., Greenbelt, MA, 1997. **Proceedings**. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 191-200.

- Maguire, G. Q. Jr.; Noz, M. E.; Lee, E. M.; Shimpf, J. H. Correlation methods for tomographic images using two e three dimensional techniques. In: Conference of Information Processing in Medical Imaging, 9., Washington, D. C., 1985.
  Proceedings. p. 266-279.
- Maguire, G. Q. J.; Noz, M. E.; and Rusinek, H.; Jaeger, J.; Kramer, E. L.; Sanger,
  J. J.; Smith, G. Graphics Applied to Medical Image Registration. IEEE
  Computers Graphics and Applications, v. 2, n. 2, p. 20-28, Mar. 1991.
- Mallat, S. G. A theory for multi-resolution signal decomposition: the wavelet representation. **IEEE Transactions on Pattern Anal. and Machine Intell.**, v. 11, n. 7, p. 674-693, July 1989a.
- Mallat, S. G. Multifrequency channel decompositions of images and wavelet models. IEEE Transactions on Acoustics, Speech, and Signal processing, v. 37, n. 12, p. 2091-2110, Dec. 1989b.
- Mallat, S. G.; Hwang, W. L. Singularity detection and processing with wavelets.

  IEEE Transactions on Information Theory, v. 38, n. 2, p. 617-643, Mar. 1992.
- Mallat, S. G.; Zhong, S. Wavelet maxima representation. In: Wavelets and Applications. Paris: Masson, 1992a. p. 207-284.
- Mallat, S. G.; Zhong, S. Characterization of Signals from multiscale edges. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 14, n. 7, p. 710-732, July 1992b.
- Mallat, S. A wavelet tour of signal processing. San Diego: Academic Press, 1998.
- Manjunath, B. S. Perceptual Grouping and Segmentation using Neural Networks. Los Angeles: University of Southern California, 1991. (USC-SIPI REPORT #188).

- Manjunath, B. S.; Chellappa, R.; Malsburg, C. A feature based approach to face recognition.
  In: IEEE Conference on Computer Vision and Pattern Recognition, Champaign, IL, 1992.
  Proceedings. Los Alamitos: IEEE Computer Society Press, 1992, p. 373-378.
- Manohar, M.; Moigne, J. L. Finding corner point correspondence from wavelet decomposition of image data. In: Image Registration Workshop, 1., Greenbelt, MA, 1997.
  Proceedings. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 243-245.
- Mascarenhas, N. D. A.; Pereira, J. A. G. Técnicas de registro de imagens por teste sequenciais de hipóteses sobre a variância. In: Simpósio Brasileiro de Sensoriamento Remoto, 2., Brasília, 1982. Anais. São José dos Campos: INPE, 1982. v. 2, p. 637-664.
- Mascarenhas, N. D. A.; Pereira, J. A. G. Image registration by sequential tests of hypotheses: gaussian and binomial techniques. In: International Symposium on Remote Sensing of Environment, 17., Ann Arbor, 1983. Proceedings. Ann Arbor: ERIM, 1983a. p. 503-512.
- Mascarenhas, N. D. A.; Percira, J. A. G. Registro de imagens por testes sequenciais de hipóteses: técnicas gaussiana e binomial. São José dos Campos: INPE, 1983b. (INPE-8384-PRE/384).
- Mascarenhas, N. D. A.; Erthal, G. J. On the relationhip between gaussian and binomial models for image registration by sequential tests of hyphoteses. In: International Congress of Photogrammetry and Remote Sensing, 15., Rio de Janeiro, 1984. Proceedings. Rio de Janeiro: ISPRS, 1984. v. 25, p. 681-689.
- Mascarenhas, N. D. A.; Erthal, G. J. Image registration by sequential tests of hypotheses: the relationship between gaussian and binomial models. **Computer** and **Graphics**, v. 16, p. 259-264, 1992.

- Mascarenhas, N. D. A., Banon, G. J. F., Candeias, A. L. B. Multispectral image data fusion under a baysian approach. International Journal Remote Sensing, v. 17, p. 1457-1471, 1996.
- Matheron, G. Eléments pour une théorie des milieux poreux. Paris: Masson, 1967.
- Melack, J. M.; Hess, L. L.; Sippel, S. Remote sensing of lakes and floodplains in the Amazon Basin. Remote Sensing Reviews, v. 10, n. 1-3, p. 127-142, 1994.
- Misiti, M.; Misiti, Y.; Oppenheim, G.; Poggi, J. Matlab, Wavelet Toolbox User's guide, New York: MathWorks, 1996.
- Moigne, J. L. Parallel registration of multi-sensor remotely sensed imagery using wavelet coefficients. In: The International Society for Optical Engineering.
   Wavelet applications. Bellingham, Wash.: SPIE, 1994. p. 432-43. (SPIE Proceedings, v. 2242)
- Moigne, J. L; Xia, W.; Chettri, S.; El-Ghazawi, T.; Kaymaz, E.; Lerner, B.; Mareboyana, M.; Netanyahu, N.; Pierce, J.; Raghavan, S.; Tilton, J. C.; Campbell, W. J.; Cromp, R. F. Towards an Intercomparison of Automated Registration Algorithms for Multiple Source Remote Sensing Data. In: Image Registration Workshop, 1., Greenbelt, MA, 1997. Proceedings. Greenbelt: NASA Goddard Space Flight Center, 1998. p. 307-315.
- Morgado, A.; Dowman, I. J. A procedure for automatic absolute orientation using aerial photographs and a map. ISPRS Journal of Photogrammetry and Remote Sensing, v. 52, n. 4, p. 169-182, Aug. 1997.
- Morlet, J.; Arens, G.; Fourgeau, E.; Giard, D. Wave Propagation and Sampling Theory. **Geophysics**, n. 2, v. 47, p. 203-236, Feb. 1982.
- Niblack, W. An introduction to digital image processing. New Jersey: Prentice/Hall, 1986.

- Oliver, C.; Quegan, S. Understanding synthetic aperture radar images.

  Norwood: Artech House, 1998.
- Persoon, E.; Fu, K. S. Shape discrimination using Fourier Descriptors. **IEEE**Transactions on Systems Man and Cybernetics, v. 7, n. 3, p. 170-179,
  Mar. 1977.
- Pratt, W. K.. Correlation techniques of image registration. IEEE Transactions on Aerospace and Electronic Systems, v. 10, n. 3, p. 353-358, May 1974.
- Pratt, W. K. **Digital image processing**. New York: John Wiley & Sons, 1991.
- Ray, B. K.; Ray, K. S. Corner detection using iterative gaussian smoothing with constant window size. Pattern Recognition, v. 28, n. 11, p. 1765-1781, Nov. 1995.
- Richards, J. R. Remote sensing digital image analysis: An Introduction. Berlin: Springer-Verlag, 1986.
- Rignot, E. J. M.; Kowk, R.; Curlander, J. C.; Pang, S. S. Automated multisensor registration: Requirements and Techniques. **Photogrammetric Engineering and Remote Sensing**, v. 57, n. 8, p. 1029-1038, Aug. 1991.
- Rioul, O.; Vetterli, M. wavelets and Signal Processing. IEEE Signal Processing Ing Magazine, v. 8, n. 4, p. 14-38, Oct. 1991.
- Safa, F.; Flouzat, G. Speckle removal on radar imagery based on mathematical morfology. Signal Processing, v. 16, n. 4, p. 319-333, Apr. 1989.
- Sant'Anna, S. J. S. Avaliação do desempenho de filtros redutores de "Speckle" em imagens de radar de abertura sintética. São José dos Campos. 210 p. (INPE-6125-TDI/586). Dissertação (Mestrado em Sensoriamento Remoto) Instituto Nacional de Pesquisas Espaciais, 1995.

- Sasakawa, K.; Isogai, F.; Ikebata, S. Personal verification system with high tolerance of poor quality fingerprints. In: The International Society for Optical Engineering. Machine vision systems integration in industry. Bellingham, Wash.: SPIE, 1990. p. 265-272. (SPIE, proceedings, v. 1386)
- Serra, J. Image analysis and mathematical morphology. London: Academic Press, 1982.
- Serra, J. Image analysis and mathematical morphology. London: Academic Press, 1988.
- Sheng, Y.; Roberge, D.; Szu, H.; Lu, T. Optical wavelet matched filters for shift-invariant pattern recognition. Optical Letters, v. 18, n. 4, p. 299-301, Feb. 1993.
- Shenk, T.; Li, J. C.; Toth, C. Towards an autonomous system for orienting digital stereopairs. **Photogrammetric Engineering and Remote Sensing**, v. 57, n. 8, p. 1057-1064, Aug. 1991.
- Strang, G. Wavelets and dilation equations: a brief Introduction. **SIAM**Review, v. 31, n. 4, p. 614-627, Dec. 1989.
- Strang, G. Wavelet Transforms versus Fourier Transforms. Bulletin of The American Mathematical Society, v. 28, n. 2, p. 288-305, Apr. 1993.
- Sun, M.; Sclabassi, R. J.; Li, C. C.; Zhang, Y.; Szu, H. H. Symetrical wavelet transforms for edge localization. Optical Engineering, v. 33, n. 7, p. 2272-2281, July 1994.
- Takeuchi, S. Image registration between SAR and TM data using DEM and slant range information. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Tokyo, 1993. Proceedings. New York: Institute of Electrical and Electronics Engineers. v. 3, p. 1351-1353.

- Teti, J. G.; Kritikos, H. N. SAR ocean image representation using wavelets. IEEE Transactions on Geoscience and Remote Sensing, v. 30, n. 5, p. 1089-1094, Sept. 1992.
- Toet, A. A morphological pyramid image decomposition. **Pattern Recognition**Letters, v. 9, n. 4, p. 255-261, May 1989.
- Ton, J.; Jain, A. K. Registering Landsat images by point matching. **IEEE**Transactions on Geoscience and Remote Sensing, v. 27, n. 5, p. 642-651,
  Sep. 1989.
- Toth, C. K.; Shenk, T. Feature-based matching for automatic image registration. ITC Journal, v. 1, n. 1, p. 40-46, 1992.
- Toutin, T. Integration of multisource data: comparison of geometric and radiometric methods. International Journal of Remote Sensing, v. 16, n. 15, p. 2795-2811, Oct. 1995.
- Trees, H. L. V. **Detection, estimation, and modulation theory**. New York: John Wiley & Sons, 1968.
- Vaidyanathan, P. P. Multirate systems and filter banks. Englewood Cliffs: Prentice Halls, 1993.
- Venkateswarlu, R.; Chatterjee, B. N. Analysis of image registration algorithms for infrared images. In: The International Society for Optical Engineering. Signal processing, sensor fusion, and target recognition. Bellingham, Wash.: SPIE, 1992. p. 442-451. (SPIE Proceedings, v. 1699)
- Ventura, A. D.; Rampini, A.; Schettini, R. Image registration by recognition of corresponding structures. IEEE Transactions on Geoscience and Remote Sensing, v. 28, n. 3, p. 305-314, May 1990.

- Vergara, O. R., Candeias, A. L. B., Kurkdjian, M. L. N. O. Data fusion in urban cartography.
  In: International Cogress for Photogrammetry and Remote Sensing, Vienna, Austria, 1996.
  Proceedings. p. 734-737.
- Wacker, A. G.; Juday, R. D.; Wolf, R. H. Jr. Landsat Image Registration- a study of system parameters. In: International Symposium on Remote Sensing of the Environment, 17., Ann Arbor, Mich., 1983. Proceedings. Ann Arbor, Mich.: Environmental Research Institute of Michigan, 1983. p. 815-825.
- Welch, R.; Jordan, T. R.; Ehlers, M. Comparative evaluations of the geodetic accuracy and cartographic potential of Landsat-4 and Landsat-5 Thematic Mapper image data. Photogrammetry Engineering Remote Sensing, v. 51, n. 9, p. 1249-1262, Sep. 1985.
- Williams, D.; Shah, M., A fast algorithm for active contour and curvature estimation. CVGIP:Image Understanding, v. 55, p. 14-26, Jan. 1992.
- Wolberg, G. **Digital image warping**. Los Alamitos: IEEE Computer Society Press, 1990.
- Wu, Y.; Maitre, H. A multiresolution approach for registration of a SPOT image and a SAR image. In: International Geoscience and Remote Sensing Symposium, Washington, D. C., 1990. Proceedings. College Park, MA: University of Maryland, 1990. v. 1, p. 635-638.
- Xu, Y.; Weaver, J. B.; Healy, D. M.; Lu, J. Wavelet transform domain filters: a spacially selective noise filtration technique. IEEE Transactions on Image Processing, v. 3, n. 6, p. 205-220, Nov. 1994.
- Zheng, Q.; Chellappa, R. A. Computational vision approach to image registration. IEEE Transactions on Image Processing, v. 2, n. 3, p. 311-326, July 1993.
- Zhou, J.; Ghosh, B. K. Image segmentation based on multiresolution filtering.
  In: IEEE Conference on Image Processing, Austin, TX, 1994. Proceedings.
  Los Alamitos: IEEE Computer Society Press, 1994. v. 3, p. 483-487.

# APÊNDICE A

### Notação

Para tornar mais fácil o entendimento da fundamentação teórica envolvida neste trabalho, uma breve revisão de alguns conceitos matemáticos usados na tese são apresentados neste apêndice.

Sejam  $A = a_i; i = 0, 1, ..., n$  e  $B = b_i; i = 0, 1, ..., n$  dois vetores de n elementos. A distância norma  $L_p$  é definida por:

$$d_1(x,y) = \sum_{i=0}^{n-1} |a_i - b_i|^p . (A.1)$$

Seja  $L^2(\mathbb{R})$  o espaço Hilbert de funções f(x) unidimensionais mensuráveis e ao quadrado integráveis .

Para  $f(x) \in L^2(\mathbb{R})$  e  $g(x) \in L^2(\mathbb{R})$ , o produto interno de f(x) com g(x) é dado por:

$$\langle g(x), f(x) \rangle = \int_{-\infty}^{+\infty} g(x)f(x)dx$$
 (A.2)

A norma (energia) de  $f(x) \in L^2(\mathbb{R})$  é definido por:

$$|| f(x) || = \int_{-\infty}^{+\infty} |f(x)|^2 dx$$
 (A.3)

Convolução de duas funções f(x) e g(x) é dado por:

$$f * g(x) = \int_{-\infty}^{+\infty} f(u)g(x - u)du . \tag{A.4}$$

A transformada de Fourier de  $f(x) \in L^2(\mathbb{R})$  denotada por F(w) é definida por

$$F(w) = \int_{-\infty}^{+\infty} f(x)e^{-iwx}dx . \tag{A.5}$$

Seja  $L^2(\mathbb{R}^2)$  o espaço Hilbert de funções bidimensionais f(x,y) mensuráveis e inte-

gráveis ao quadrado.

A norma (energia) de  $f(x,y) \in L^2(\mathbb{R}^2)$  é escrito por:

$$|| f(x,y) || = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |f(x,y)|^2 dx dy$$
 (A.6)

A transformada de Fourier de  $f(x,y) \in L^2(\mathbb{R}^2)$  denotada por  $F(w_x,w_y)$  é definida por

$$F(w_x, w_y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) e^{-i(w_x x + w_y y)} dx dy . \tag{A.7}$$

Teorema de Parseval:

Se a transformada de Fourier de g(x,y) é igual a  $G(w_x,w_y)$ , então

$$\int \int_{-\infty}^{\infty} || g(x,y) ||^2 dx dy = \int \int_{-\infty}^{\infty} || G(w_x, w_y) ||^2 dw_x dw_y.$$
 (A.8)

# APÊNDICE B

#### Prova de teoremas

Neste apêndice são apresentadas as provas de dois teoremas (Equações 3.17 e 3.21) mencionados no Capítulo 3 e que não foram demonstrados nas referências em que foram citados.

### PROVA 1

Vamos provar que:

$$\lim_{w\to 0}\Phi(w)=1.$$

dado que

$$\sum_{j=-\infty}^{+\infty} \Psi(2^j w) \mathcal{X}(2^j w) = 1 , \qquad (B.1)$$

e

$$\sum_{j=-\infty}^{+\infty} \Psi(2^j w) \mathcal{X}(2^j w) = 1.$$
(B.2)

### Prova:

A Equação B.1 pode ser escrita da seguinte forma

$$\sum_{j=-\infty}^{0} \Psi(2^{j}w)\mathcal{X}(2^{j}w) + \sum_{j=1}^{+\infty} \Psi(2^{j}w)\mathcal{X}(2^{j}w) = 1.$$
 (B.3)

Usando a Equação B.2, podemos reescrever a Equação B.3

$$\sum_{j=-\infty}^{0} \Psi(2^{j}w)\mathcal{X}(2^{j}w) + |\Phi(w)|^{2} = 1,$$

$$|\Phi(w)|^{2} = 1 - \sum_{j=-\infty}^{0} \Psi(2^{j}w)\mathcal{X}(2^{j}w). \tag{B.4}$$

Dado que  $\Psi(0) = 0$  (filtro passa-banda), de B.4 temos que

$$\lim_{w\to 0} |\Phi(w)|^2 = 1$$
.

### PROVA 2

Vamos provar que:

$$||A_1 f(x)||^2 - ||A_2 f||^2 = \sum_{j=1}^L W_{2^j} f(x) * \xi_{2^j}(x) .$$

dado que temos:

$$S_1 f(w) = F(w) \cdot \Phi(w)$$

$$S_{2^L} f(w) = F(w) \cdot \Phi(2^L w) . \tag{B.5}$$

$$|\Phi(w)|^2 = \sum_{j=1}^{L} \Psi(2^j w) \mathcal{X}(2^j w) + |\Phi(2^L w)|^2$$
 (B.6)

$$\Omega_{2^j}(w) = F(w) \cdot \Psi(2^j w) . \tag{B.7}$$

Prova:

Seja F(w) a transformada de Fourier de um sinal f(x). Multiplicando a Equação B.6 por F(w), temos:

$$|F(w)\Phi(w)|^2 = \sum_{j=1}^{L} F(w)\Psi(2^j w) | + |F(w)\Phi(2^L w)|^2$$
 (B.8)

Usando as Equações B.5 e B.7 em B.8 temos:

$$\mid \mathcal{S}_{1}f(w)\mid^{2} = \sum_{j=1}^{L}\Omega_{2^{j}}(w)\mathcal{X}(2^{j}w) + \mid \mathcal{S}_{2^{L}}f(w)\mid^{2}$$

$$|\mathcal{S}_1 f(w)|^2 - |\mathcal{S}_{2^L} f(w)|^2 = \sum_{j=1}^L \Omega_{2^j}(w)|.$$
 (B.9)

Pelo teorema de Parseval, da Equação B.9 temos:

$$||A_1 f(x)||^2 - ||A'_{2L} f(x)||^2 = \sum_{j=1}^{L} W_{2j} f(x) * \xi_{2j}(x).$$

c.q.d.

### APÊNDICE C

### Refinamento dos parâmetros da transformação

No processo de registro as imagens são decompostas em L níveis de resolução. O processamento é realizado primeiro no nível mais alto (L), indo para os níveis progressivamente menores. Em cada nível de processamento da pirâmide j-1, os parâmetros da função de transformação são atualizados usando os parâmetros obtidos no nível anterior j e os parâmetros obtidos no nível corrente j-1 durante o processo de refinamento do registro. As Equações 6.9 e 6.10 mostram, respectivamente, como os parâmetros das funções de transformação do tipo similaridade e afim, no nível corrente j-1, podem ser obtidos. O desenvolvimento para a obtenção destas equações é mostrado neste apêndice.

Sejam  $\mathcal{T}_j$ ,  $\mathcal{T}_{j-1}$  e  $\mathcal{T}_{j-1}^r$  as funções de transformação obtidas nos níveis de processamento j e j-1 e no processo de refinamento no nível de processamento j-1, respectivamente.

Vamos considerar que a função de transformação  $\mathcal{T}$ , seja do tipo similaridade. Sejam  $\{s_j, \theta_j, \Delta x_j, \Delta y_j\}$ ,  $\{s_{j-1}, \theta_{j-1}, \Delta x_{j-1}, \Delta y_{j-1}\}$  e  $\{s_{j-1}^r, \theta_{j-1}^r, \Delta x_{j-1}^r, \Delta y_{j-1}^r\}$  os parâmetros das funções de transformação  $\mathcal{T}_j$ ,  $\mathcal{T}_{j-1}$  e  $\mathcal{T}_{j-1}^r$ , respectivamente. Os parâmetros da transformação total  $\mathcal{T}_{j-1}$ , no nível j-1, são obtidos pela combinação das transformações  $\mathcal{T}_{j-1}^r$  e  $\mathcal{T}_j$ :

$$\mathcal{T}_{j-1} = \mathcal{T}_{j-1}^{\tau} \circ \mathcal{T}_{j}, \tag{C.1}$$

onde o representa a composição de funções.

As transformações  $\mathcal{T}_{j-1}^r$  e  $\mathcal{T}_j$  que levam  $(x_{j-1}, y_{j-1})$  para  $\mathcal{T}_{j-1}^r(x_{j-1}, y_{j-1})$  e  $(x_j, y_j)$  para  $\mathcal{T}_j(x_j, y_j)$  são definidas, respectivamente, por :

$$\mathcal{T}_{j-1}^{r} \begin{pmatrix} x_{j-1} \\ y_{j-1} \end{pmatrix} = s_{j-1}^{r} \begin{pmatrix} \cos(\theta_{j-1}^{r}) & \sin(\theta_{j-1}^{r}) \\ -\sin(\theta_{j-1}^{r}) & \cos(\theta_{j-1}^{r}) \end{pmatrix} \begin{pmatrix} x_{j-1} \\ y_{j-1} \end{pmatrix} + \begin{pmatrix} \Delta x_{j-1}^{r} \\ \Delta y_{j-1}^{r} \end{pmatrix} \tag{C.2}$$

 $\mathbf{e}$ 

$$\mathcal{T}_{j} \begin{pmatrix} x_{j} \\ y_{j} \end{pmatrix} = s_{j} \begin{pmatrix} \cos(\theta_{j}) & \sin(\theta_{j}) \\ -\sin(\theta_{j}) & \cos(\theta_{j}) \end{pmatrix} \begin{pmatrix} x_{j} \\ y_{j} \end{pmatrix} + \begin{pmatrix} 2\Delta x_{j}^{r} \\ 2\Delta y_{j}^{r} \end{pmatrix}$$
(C.3)

Para obter se  $\mathcal{T}_{j-1}$ , substitui-se as Equações C.3 e C.2 em C.1:

$$\mathcal{T}_{j-1}\begin{pmatrix} x_{j-1} \\ y_{j-1} \end{pmatrix} = s_{j-1}^{r} s_{j} \begin{pmatrix} \cos(\theta_{j-1}^{r}) & \sin(\theta_{j-1}^{r}) \\ -\sin(\theta_{j-1}^{r}) & \cos(\theta_{j-1}^{r}) \end{pmatrix} \begin{pmatrix} \cos(\theta_{j}) & \sin(\theta_{j}) \\ -\sin(\theta_{j}) & \cos(\theta_{j}) \end{pmatrix} \begin{pmatrix} x_{j} \\ y_{j} \end{pmatrix} + s_{j-1}^{r} \begin{pmatrix} \cos(\theta_{j-1}^{r}) & \sin(\theta_{j-1}^{r}) \\ -\sin(\theta_{j-1}^{r}) & \cos(\theta_{j-1}^{r}) \end{pmatrix} \begin{pmatrix} 2\Delta x_{j} \\ 2\Delta y_{j} \end{pmatrix} + \begin{pmatrix} \Delta x_{j-1}^{r} \\ \Delta y_{j-1}^{r} \end{pmatrix}$$
 (C.4)

Continuando o desenvolvimento de C.4:

$$\mathcal{T}_{j-1} \begin{pmatrix} x_{j-1} \\ y_{j-1} \end{pmatrix} = s_{j-1}^{r} s_{j} \begin{pmatrix} \cos(\theta_{j-1}^{r} + \theta_{j}) & \sin(\theta_{j-1}^{r} + \theta_{j}) \\ -\sin(\theta_{j-1}^{r} + \theta_{j}) & \cos(\theta_{j-1}^{r} + \theta_{j}) \end{pmatrix} \begin{pmatrix} x_{j} \\ y_{j} \end{pmatrix} + \begin{pmatrix} 2s_{j-1}^{r} (\Delta x_{j} \cos(\theta_{j-1}^{r}) + \Delta y_{j} \sin(\theta_{j-1}^{r}) + \Delta x_{j-1}^{r} \\ 2s_{j-1}^{r} (-\Delta x_{j} \sin(\theta_{j-1}^{r}) + \Delta y_{j} \cos(\theta_{j-1}^{r}) + \Delta y_{j-1}^{r} \end{pmatrix}.$$
(C.5)

Desta forma os parâmetros da função de similaridade no nível j-1 são:

$$\begin{pmatrix} s_{j-1} \\ \theta_{j-1} \\ \Delta x_{j-1} \\ \Delta y_{j-1} \end{pmatrix} = \begin{pmatrix} s_{j-1}^{\mathsf{r}} s_{j} \\ \theta_{j-1}^{\mathsf{r}} + \theta_{j} \\ 2s_{j-1}^{\mathsf{r}} (\cos(\theta_{j-1}^{\mathsf{r}}) \Delta x_{j} + \sin(\theta_{j-1}^{\mathsf{r}}) \Delta y_{j}) + \Delta x_{j-1}^{\mathsf{r}} \\ 2s_{j-1}^{\mathsf{r}} (\cos(\theta_{j-1}^{\mathsf{r}}) \Delta x_{j} + \sin(\theta_{j-1}^{\mathsf{r}}) \Delta y_{j}) + \Delta y_{j-1}^{\mathsf{r}} \end{pmatrix}. \quad (C.6)$$

c.q.d.

Os parâmetros da função afim podem ser obtidos seguindo o mesmo procedimento descrito acima.