

NEW TECHNIQUES FOR THE AUTOMATIC REGISTRATION OF MICROWAVE AND OPTICAL REMOTELY SENSED IMAGES

by

Paul Maxwell DARE

**Submitted for the degree of
Doctor of Philosophy
at the
University of London**



**Department of Geomatic Engineering
University College London
Gower Street
London WC1E 6BT**

April 1999

ProQuest Number: 10609146

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10609146

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

Remote sensing is a remarkable tool for monitoring and mapping the land and ocean surfaces of the Earth. Recently, with the launch of many new Earth observation satellites, there has been an increase in the amount of data that is being acquired, and the potential for mapping is greater than ever before. Furthermore, sensors which are currently operational are acquiring data in many different parts of the electromagnetic spectrum. It has long been known that by combining images that have been acquired at different wavelengths, or at different times, the ability to detect and recognise features on the ground is greatly increased. This thesis investigates the possibilities for automatically combining radar and optical remotely sensed images.

The process of combining images, known as data integration, is a two step procedure: geometric integration (image registration) and radiometric integration (data fusion). Data fusion is essentially an automatic procedure, but the problems associated with automatic registration of multisource images have not, in general, been resolved. This thesis proposes a method of automatic image registration based on the extraction and matching of common features which are visible in both images. The first stage of the registration procedure uses patches as the matching primitives in order to determine the approximate alignment of the images. The second stage refines the registration results by matching edge features. Throughout the development of the proposed registration algorithm, reliability, robustness and automation were always considered priorities. Tests with both small images (512×512 pixels) and full scene images showed that the algorithm could successfully register images to an acceptable level of accuracy.

TABLE OF CONTENTS

<i>Abstract</i>	2
<i>Table of contents</i>	3
<i>List of figures</i>	7
<i>List of tables</i>	11
Chapter 1 Introduction	13
1.1 Data integration	13
1.2 Image registration	15
1.3 Data fusion	17
1.4 Applications of multi-image data sets	20
1.5 Problems associated with data integration	24
1.6 Automatic multi-image data integration	25
1.7 Discussion	26
1.8 Glossary of terms	27
Chapter 2 Image registration	32
2.1 Introduction	32
2.2 Rigorous geometric correction of images	32
2.2.1 Introduction	32
2.2.2 Orthorectification of optical images	32
2.2.3 Geocoding of SAR images	35
2.2.4 Conclusions	36
2.3 Two dimensional image registration	36
2.3.1 Introduction	36
2.3.2 Tie point selection	37
2.3.3 Transformation functions	37
2.3.4 Calculation of transformation parameters	45
2.3.5 Resampling	46
2.3.6 Conclusions	47
2.4 Automatic image registration	47
2.4.1 Introduction	47
2.4.2 Extraction of primitives	48
2.4.3 Matching of primitives	49
2.4.4 Determination of the transformation function	49
2.4.5 Review of automatic registration techniques	49
2.4.6 Difficulties associated with automatic image registration	54
2.4.7 Conclusions	57
2.5 Proposal of this thesis	59
2.5.1 Introduction	59

2.5.2	The "ideal" automatic image registration system	59
2.5.3	Elements of automatic registration researched in detail	62
2.5.4	Conclusions	64
Chapter 3 Data acquisition and image selection		66
3.1	Introduction	66
3.2	Acquisition and properties of SAR data	66
3.2.1	Introduction	66
3.2.2	Resolution	69
3.2.3	Geometric distortions	73
3.2.4	Radiometric attributes	73
3.2.5	The ERS family of satellites	77
3.3	Acquisition and properties of SPOT data	80
3.3.1	Introduction	80
3.3.2	The pushbroom sensor	80
3.3.3	Image characteristics	81
3.3.4	The SPOT satellite series	83
3.4	Image selection	87
3.4.1	Introduction	87
3.4.2	SPOT data	88
3.4.3	SAR data	89
3.4.4	Selected test sites	90
3.4.5	Conclusions	92
3.5	Discussion	93
Chapter 4 Radiometric preprocessing		96
4.1	Introduction	96
4.2	Preprocessing of SAR data	96
4.2.1	Introduction	96
4.2.2	Simple speckle reduction filters	97
4.2.3	Adaptive, rigorous filters	100
4.2.4	Adaptive, non-rigorous filters	103
4.2.5	Improved traditional filters	106
4.2.6	Comparison of speckle reduction filters	108
4.2.7	Conclusions	113
4.3	Preprocessing of SPOT data	114
4.3.1	Introduction	114
4.3.2	Smoothing of SPOT images	114
4.3.3	Conclusions	116
4.4	Discussion	116
Chapter 5 Feature extraction		118
5.1	Introduction	118
5.2	Edge extraction	120
5.2.1	Introduction	120
5.2.2	Gradient operators for edge detection	121
5.2.3	The SCANEDGE algorithm	125
5.2.4	Other methods of linear feature extraction	127
5.2.5	Post-processing of edge images	129
5.2.6	Conclusions	131
5.3	Patch extraction	132
5.3.1	Introduction	132
5.3.2	Thresholding	133
5.3.3	Homogeneous patch extraction	144
5.3.4	Segmentation	146
5.3.5	Boundary to region conversion	155

5.3.6	Conclusions	156
5.4	Post-processing of extracted polygons	158
5.4.1	Introduction	158
5.4.2	Unique grey value assignment	158
5.4.3	Clutter removal	159
5.4.4	Polygon to edge conversion	161
5.4.5	Conclusions	162
5.5	Discussion	162
Chapter 6	Patch Matching	165
6.1	Introduction	165
6.2	Extraction of patches from SAR and SPOT images	166
6.2.1	Introduction	166
6.2.2	Thresholding	166
6.2.3	Homogeneous patch extraction	170
6.2.4	Segmentation	174
6.2.5	Conclusions	182
6.3	Patch matching	183
6.3.1	Introduction	183
6.3.2	Patch attribute encoding	184
6.3.3	Matching	186
6.3.4	Image registration	194
6.3.5	Conclusions	196
6.4	New techniques for improved patch matching	197
6.4.1	Introduction	197
6.4.2	Automatic parameter selection	198
6.4.3	Combined use of multiple feature extraction results	201
6.4.4	Integrated feature extraction and matching	206
6.4.5	Results of refined patch matching	208
6.4.6	Conclusions	213
6.5	Discussion	214
Chapter 7	Improving automatic registration using edge matching	217
7.1	Introduction	217
7.2	Extraction of linear features from SAR and SPOT images	220
7.2.1	Introduction	220
7.2.2	First derivative edge detection	221
7.2.3	Second derivative edge detection	226
7.2.4	CAESAR edge detection	228
7.2.5	Conclusions	230
7.3	Edge matching	231
7.3.1	Introduction	231
7.3.2	Edge matching by dynamic programming	232
7.3.3	Edge matching results	233
7.3.4	Image registration	235
7.3.5	Conclusions	239
7.4	Discussion	240
Chapter 8	Automatic registration of full scene images	242
8.1	Introduction	242
8.2	Pre-registration processing	244
8.2.1	Introduction	244
8.2.2	Initial alignment	244
8.2.3	Creation of image tiles	247
8.2.4	Conclusions	248

8.3	Automatic registration using feature matching	248
8.3.1	Introduction	248
8.3.2	Polygon extraction and matching	249
8.3.3	Image registration	253
8.3.4	Refinement using edge matching	261
8.3.5	Conclusions	266
8.4	Automatic orthorectification of SPOT data	267
8.5	Discussion	269
<i>Chapter 9 Discussion</i>		272
<i>Chapter 10 References</i>		278

LIST OF FIGURES

Figure 1.1 Data integration processing chain	14
Figure 1.2 The electromagnetic spectrum and atmospheric absorption	15
Figure 1.3 (a) Image registration	16
Figure 1.3 (b) Image rectification	16
Figure 1.4 NDVI image of Europe for September 1996 (Gredel and Schröter, 1997)	18
Figure 1.5 (a) Three channels of SPOT multispectral data	19
Figure 1.5 (b) Correctly aligned SPOT XS channels (left), misaligned channels (right)	19
Figure 1.6 (a) SIR-C/X-SAR images of Manaus, 12/04/94 & 3/10/94 respectively (© NASA)	20
Figure 1.6 (b) Difference image representing changes in flooding (© NASA)	21
Figure 1.6 (c) Inundation map of Manaus, Brazil (© NASA)	21
Figure 1.7 (a) L-band, HV polarization (© NASA)	21
Figure 1.7 (b) C-band, HV polarization (© NASA)	22
Figure 1.7 (c) X-band, VV polarization (© NASA)	22
Figure 1.8 False colour composite of Rodonia, western Brazil (© NASA)	23
Figure 1.9 (a) 5.8m panchromatic and 23m multispectral IRS-1C images (© Euromap)	23
Figure 1.9 (b) 5.8m panchromatic image with registered 23m multispectral image	24
Figure 1.9 (c) IHS merged image	24
Figure 2.1 Distortion due to relief in aerial photographs	34
Figure 2.2 Orthoimage production	34
Figure 2.3 Manual image registration processing chain	37
Figure 2.4 Space Shuttle handheld camera images of Lake Kariba (© NASA)	40
Figure 2.5 Mosaic created using affine transformation	40
Figure 2.6 Mosaic created using quadratic transformation	41
Figure 2.7 Non-linear function (red curve) with 4 th and 8 th order best fit polynomials (green and blue curves respectively)	42
Figure 2.8 Resampling schemes	46
Figure 2.9 Tie point extraction from matched polygons	49
Figure 2.10 Radar and optical images of Death Valley, USA (© NASA)	55
Figure 2.11 Radar and optical images of Southern France	56
Figure 2.12 Registered radar and optical images of Southern France	57
Figure 2.13 The ideal automatic image registration system	62
Figure 2.14 Flow chart of proposed image registration procedure based on patch and edge matching	63
Figure 3.1 The SLAR sensor	69
Figure 3.2 Geometry of the SLAR sensor	70
Figure 3.3 Azimuth resolution of the SLAR sensor (after Lillesand and Kiefer, 1994)	71
Figure 3.4 SAR and beamwidth	72
Figure 3.5 Distortions due to relief in optical and radar images (after Lillesand and Kiefer, 1994)	73
Figure 3.6 Mean and standard deviation of pixels in ERS-1 PRI SAR image	75
Figure 3.7 Sidelighting and shadow in SAR images (Lillesand and Kiefer, 1994)	76
Figure 3.8 The ERS-1 satellite (© ESA)	77
Figure 3.9 ERS orbital geometry (© ESA)	78
Figure 3.10 The Envisat platform (left) and ERS platform (right) (© ESA)	79
Figure 3.11 The pushbroom scanner	81
Figure 3.12 Cross section of the HRV sensor (© CNES)	84
Figure 3.13 HRV spectral bands (© CNES)	85
Figure 3.14 HRV viewing geometry (© CNES)	85
Figure 3.15 SPOT stereo viewing capability (© CNES)	86
Figure 3.16 Study area in Southern France	88
Figure 3.17 Reduced size full scene SPOT image	89
Figure 3.18 Example reduced size full scene SAR image (10:28:48)	90

Figure 3.19 Beaucaire test images	91
Figure 3.20 Camargue test images	91
Figure 3.21 Entressen test images	91
Figure 3.22 Istres test images	92
Figure 4.1 Mean filter applied to Beaucaire image (left: 3×3, right: 9×9)	99
Figure 4.2 Plus shaped mean filter applied to Beaucaire image (left: 3×3, right: 9×9)	99
Figure 4.3 Median filter applied to Beaucaire test image (left: 3×3, right: 9×9)	100
Figure 4.4 Gaussian smoothing of Beaucaire test image (left: 3×3, right: 9×9)	100
Figure 4.5 Frost filter (left) and Lee filter (right) applied to Beaucaire test image	102
Figure 4.6 Kuan filter (left) and MAP filter (right) applied to Beaucaire test image	103
Figure 4.7 Nine masks used in MHN filter	104
Figure 4.8 Sigma filter, MHN filter, and modified k-average filter applied to Beaucaire test image	105
Figure 4.9 Sigma filter and sigma-median filter applied to Beaucaire test image	107
Figure 4.10 Traditional MHN filter and MHN-median filter applied to Beaucaire test image	107
Figure 4.11 k-average filter and the k-median filter applied to Beaucaire test image	108
Figure 4.12 Location of step edges in Beaucaire image	109
Figure 4.13 Perspective view of step edge in Beaucaire test image	110
Figure 4.14 Perspective view of step edge after application of mean filter (left) and median filter (right)	112
Figure 4.15 Perspective view of step edge after application of MHN filter (left) and MHN-median filter (right)	112
Figure 4.16 Perspective view of step edge after application of sigma-median filter (left) and k-median filter (right)	113
Figure 4.17 Smoothing algorithms applied to SPOT image of Beaucaire	115
Figure 5.1 Original and edge strength images generated using Sobel operator (© Playboy Inc.)	122
Figure 5.2 Step edge with first and second order derivatives	123
Figure 5.3 Zero crossings in test image, w=5 (left) and w=9 (right)	124
Figure 5.4 Edges extracted from Entressen images using SCANEDGE with three different values of p	126
Figure 5.5 Four processing steps in Hough transform procedure	128
Figure 5.6 SPOT image of Istres (left) and Sobel edge strength image(right)	130
Figure 5.7 Non-maximal suppression proposed by Lewis (1988)	130
Figure 5.8 Weak edges suppressed using non-maximal suppression	131
Figure 5.9 Thresholded edge image	131
Figure 5.10 Original greyscale image (left) and image thresholded at grey level of 112 (right) (© Playboy Inc.)	134
Figure 5.11 Single feature image (left) and corresponding bimodal histogram (right)	135
Figure 5.12 Radar image of oil slick (left) and image thresholded at grey level of 100 (right)	136
Figure 5.13 Original histogram (a), and histograms smoothed with 3, 5, 7 grey level kernels (b, c, d respectively).	137
Figure 5.14 Ten iterations of histogram smoothing using 7 grey level kernel	138
Figure 5.15 Oil slick image (left) and automatically thresholded image (right)	140
Figure 5.16 SAR image of river and town (left) and corresponding histogram (right)	140
Figure 5.17 Histograms of smoothed Beaucaire image	141
Figure 5.18 Automatically thresholded SAR image of region near Beaucaire	142
Figure 5.19 SPOT image of agricultural land near Entressen (left) and corresponding histogram (right)	142
Figure 5.20 SPOT image of agricultural land thresholded at grey level 119	143
Figure 5.21 SPOT image of Istres (left) and extracted homogeneous patches (right)	145
Figure 5.22 SAR image of Istres (left) and extracted homogeneous patches (right)	145
Figure 5.23 SAR image segmented with MUM using four different values of p.	148
Figure 5.24 SAR image segmented with RWSEG using two different values of e and j.	148
Figure 5.25 SPOT image segmented with REGSEG using four different values of d (2, 3, 4 and 5).	152
Figure 5.26 SPOT image segmented with OPTISEG using different values of T, μ and ν .	153
Figure 5.27 Edge map of SPOT image of Istres	155
Figure 5.28 SPOT image segmented without edge map (left) and with edge map (right)	155
Figure 5.29 Clutter in automatically thresholded image of an oil slick	160
Figure 5.30 Examples of 5×5 and a 7×7 windows used for clutter removal	160
Figure 5.31 Thresholded oil slick image with clutter removed using RM_CLUT (left) and RM_SMALL_PATCHES (right).	161
Figure 5.32 Four possible positions of edge extraction kernel in relation to an edge	162
Figure 6.1 SAR image of Istres and corresponding thresholded image (threshold = 81)	167
Figure 6.2 SPOT image of Istres and corresponding thresholded image (threshold = 156)	167

Figure 6.3 SAR image of Camargue and corresponding thresholded image (threshold = 66)	168
Figure 6.4 SPOT image of Camargue and corresponding thresholded image (threshold = 68)	168
Figure 6.5 SAR and SPOT images of Istres after clutter removal	169
Figure 6.6 SAR and SPOT images of Camargue after clutter removal	169
Figure 6.7 Flow chart summarizing automatic thresholding	170
Figure 6.8 Homogeneous patch extraction results for SAR image of Istres	171
Figure 6.9 Homogeneous patch extraction results for SPOT image of Istres	171
Figure 6.10 Homogeneous patch extraction results for SAR image of Camargue	171
Figure 6.11 Homogeneous patch extraction results for SPOT image of Camargue	172
Figure 6.12 Post-processed patches extracted using homogeneous patch extraction for SAR (left) and SPOT (right) images of Istres	172
Figure 6.13 Post-processed patches extracted using homogeneous patch extraction for SAR (left) and SPOT (right) images of Camargue	173
Figure 6.14 Flow chart summarizing homogeneous patch extraction	173
Figure 6.15 Istres SAR image segmented using MUM algorithm with $p = 6$	174
Figure 6.16 Camargue SAR image segmented using MUM algorithm with $p = 6$	174
Figure 6.17 Patches extracted from Istres (left) and Camargue (right) using MUM segmentation	175
Figure 6.18 Istres SAR image segmented using RWSEG with $e = 3.29$ and $j = 6$	176
Figure 6.19 Camargue SAR image segmented using RWSEG with $e = 3.29$ and $j = 4$	176
Figure 6.20 Patches extracted from Istres (left) and Camargue (right) using RWSEG segmentation	177
Figure 6.21 Flow charts summarizing MUM and RWSEG algorithms	178
Figure 6.22 Istres SPOT image segmented using REGSEG with $d = 3$	179
Figure 6.23 Camargue SPOT image segmented using REGSEG with $d = 3$	179
Figure 6.24 Patches extracted from Istres (left) and Camargue (right) using REGSEG segmentation	180
Figure 6.25 Istres SPOT image segmented using OPTISEG with $T = 5$, $\mu = 0.05$, $\nu = -0.2$	180
Figure 6.26 Camargue SPOT image segmented using OPTISEG with $T = 5$, $\mu = 0.2$, $\nu = -0.2$	181
Figure 6.27 Patches extracted from Istres (left) and Camargue (right) using OPTISEG segmentation	181
Figure 6.28 Flow chart of optical segmentation procedure	182
Figure 6.29 Area, perimeter and bounding rectangle of typical patch	184
Figure 6.30 Patches from SPOT image of Istres with corresponding reference numbers	186
Figure 6.31 Flow chart summarizing patch matching procedure	187
Figure 6.32 Matched patches from SAR and SPOT images of Istres	192
Figure 6.33 Matched patches extracted from SPOT (left) and SAR (right) images of Camargue	195
Figure 6.34 Matched patches extracted from SPOT (left) and SAR (right) images of Istres	195
Figure 6.35 Best matched patches from SAR and SPOT images of Camargue (left and right respectively)	203
Figure 6.36 Best matched patches from SAR and SPOT images of Istres (left and right respectively)	204
Figure 6.37 Best matched patches from SAR (left) and SPOT (right) images of Camargue	205
Figure 6.38 Best matched patches from SAR (left) and SPOT (right) images of Istres	206
Figure 6.39 Flow chart summarizing integrated feature extraction and matching	207
Figure 6.40 Flow chart summarizing image matching with refinements	209
Figure 6.41 Final matched patches from SAR (left) and SPOT (right) images of Camargue	210
Figure 6.42 Final matched patches from SAR (left) and SPOT (right) images of Istres	210
Figure 6.43 Results of patch based registration for Camargue (left) and Istres (right)	212
Figure 7.1 Selection of edges for matching using masks	219
Figure 7.2 Edge images generated from Sobel edge detector applied to smoothed Camargue SPOT images	222
Figure 7.3 Edge images generated from Sobel edge detector applied to smoothed Istres SPOT images	222
Figure 7.4 Edge images generated from Sobel edge detector applied to Camargue SAR images	224
Figure 7.5 Edge images generated from Sobel edge detector applied to Istres SAR images	225
Figure 7.6 Flow chart summarizing linear feature extraction using first derivative edge detection	226
Figure 7.7 LoG algorithm applied to SAR and SPOT images of Camargue	227
Figure 7.8 LoG algorithm applied to SAR and SPOT images of Istres	228
Figure 7.9 Flow chart summarizing the application of the LoG algorithm	228
Figure 7.10 Camargue image and result of CAESAR edge detection	229
Figure 7.11 Istres image and result of CAESAR edge detection	229
Figure 7.12 Flow chart summarizing the application of the CAESAR algorithm	230
Figure 7.13 Linear feature matching	232
Figure 7.14 Distribution of residuals for two Camargue matching results	237
Figure 7.15 Distribution of residuals for two Istres matching results	237
Figure 7.16 Outlines of transformed SPOT features (green) overlaid on SAR image	239
Figure 8.1 Approximate registration using ephemeris data (orthographic ground co-ordinate system)	245

<i>Figure 8.2 Approximate registration using four manually selected tie points (SAR image co-ordinate system)</i>	246
<i>Figure 8.3 SAR image split into tiles</i>	247
<i>Figure 8.4 SPOT image split into tiles</i>	248
<i>Figure 8.5 Distribution of tiles selected to be matched</i>	250
<i>Figure 8.6 Number of correct matches found for each processed tile</i>	251
<i>Figure 8.7 Matched features from SAR image</i>	252
<i>Figure 8.8 Matched features from SPOT image</i>	252
<i>Figure 8.9 Visual result of automatic registration using patch matching</i>	253
<i>Figure 8.10 Two sets of residuals overlaid on original SAR image</i>	256
<i>Figure 8.11 Matched patches from tile at (1,3)</i>	257
<i>Figure 8.12 Matched patches from tile at (1,8)</i>	257
<i>Figure 8.13 Matched patches from tile at (2,3)</i>	258
<i>Figure 8.14 Matched patches from tile at (2,8)</i>	258
<i>Figure 8.15 Matched patches from tile at (4,4)</i>	258
<i>Figure 8.16 Matched patches from tile at (6,2)</i>	259
<i>Figure 8.17 Matched patches from tile at (6,5)</i>	259
<i>Figure 8.18 Matched patches from tile at (7,11)</i>	259
<i>Figure 8.19 Matched patches from tile at (8,6)</i>	260
<i>Figure 8.20 Matched patches from tile at (8,9)</i>	260
<i>Figure 8.21 Spatial distribution of tie points found using edge matching</i>	263
<i>Figure 8.22 Refined residuals from edge matching overlaid onto SAR image</i>	264
<i>Figure 8.23 Visual result of automatic registration refined using edge matching</i>	265
<i>Figure 8.24 Method of automatic registration using photogrammetric techniques</i>	268

LIST OF TABLES

Table 1.1 Definitions of some terms used in this thesis	31
Table 2.1 Summary of automatic registration techniques	54
Table 3.1 Dependence of σ_v on the number of looks, N (Lee, 1986).	75
Table 3.2 Characteristics of ERS orbit and AMI specifications	78
Table 3.3 Specifications of Envisat's ASAR instrument	80
Table 3.4 SPOT orbital parameters and sensor characteristics	83
Table 3.5 Summary of SPOT images used in this study	88
Table 3.6 Summary of SAR images used in this study	89
Table 4.1 Speckle indices and edge retention indicators for various speckle reduction filters	111
Table 5.1 Number of turning points for different combinations of kernels and iterations	138
Table 5.2 Threshold levels for different combinations of kernels and iterations	139
Table 5.3 Summary of feature extraction algorithms	163
Table 6.1 Attributes of patches in SPOT image of Istres processed with homogeneous patch extraction (all units are pixels)	186
Table 6.2 Example of cost function matching results	189
Table 6.3 Refined matching results with correct matches highlighted	191
Table 6.4 Final matching results	192
Table 6.5 Summary of all matching results for Istres and Camargue	193
Table 6.6 Co-ordinates of Camargue and Istres tie points with residuals	196
Table 6.7 Number of correct matches for SPOT homogeneous patch extraction and SAR homogeneous patch extraction	199
Table 6.8 Number of correct matches for SPOT REGSEG and SAR homogeneous patch extraction	199
Table 6.9 Number of correct matches for SPOT homogeneous patch extraction and SAR MUM	200
Table 6.10 Number of correct matches for SPOT REGSEG and SAR MUM	200
Table 6.11 Matched patches from Camargue images	202
Table 6.12 Best matched patches from Camargue images	203
Table 6.13 Best matching results for Camargue image	205
Table 6.14 Best matching results for Istres image	205
Table 6.15 Results of combined feature extraction and matching	208
Table 6.16 Details of matched patches for Camargue images	210
Table 6.17 Details of matched patches for Istres images	211
Table 6.18 Co-ordinates of Camargue tie points with residuals	211
Table 6.19 Co-ordinates of Istres tie points with residuals	211
Table 6.20 Coefficients of affine transformation function for Camargue and Istres	212
Table 7.1 Number of pixels extracted for each edge extraction algorithm	231
Table 7.2 Number of tie points found with percentages for Camargue image using different combinations of edge extraction algorithms	234
Table 7.3 Number of tie points found with percentages for Istres image using different combinations of edge extraction algorithms	234
Table 7.4 RMS residuals for Camargue image transformation for different combinations of edge extraction algorithms	235
Table 7.5 RMS residuals for Istres image transformation edge different combinations of edge extraction algorithms	236
Table 7.6 Coefficients of affine transformation function for Camargue and Istres	238
Table 8.1 Residuals in x and y for two sets of check points	254
Table 8.2 RMS residuals for two sets of check points	255
Table 8.4 Number of tie points found and RMS residual for each processed tile	262
Table 8.5 RMS residuals calculated using tie points from edge matching	263

To Sharon and Bradley

Chapter 1 INTRODUCTION

1.1 Data integration

Data integration is fast becoming a cornerstone of remote sensing. The ability to accurately recognise features on the Earth's surface relies heavily on the quality and the nature of the images used to map those features. Although Earth observation sensors are becoming ever more sensitive, it is not always sufficient to use only one type of sensor to map a particular area on the ground; if all the features in that area are to be mapped accurately, then images from different sensors may be required. The process of combining images from different sensors in this way is known as data integration. A consequence of the integration of data sets from multiple sensors is increased spatial, spectral and temporal resolution, increased reliability, and reduced ambiguity (Keys et al., 1990; Rogers and Wood, 1990).

Nowadays, with the number of Earth observation satellites increasing at a phenomenal rate, the acquisition of images from a multitude of sensors is not a significant problem. Furthermore, developments in geographical information systems (GIS) mean that spatial and temporal analysis of images is becoming far easier. However, a bottleneck in the processing chain still exists between the acquisition and interpretation of multiple images – this is the geometric and radiometric processing of images which needs to take place so they can be compared and analysed in a common reference frame (Dowman, 1998).

Data integration describes the complete process of merging images from different sources in order to create a single data set that can be used for analysis of features in the merged image. Data integration does not only refer to combining data from different sensors – it also covers integrating data acquired by the same sensor, but at different times, in other words, temporal data integration as opposed to multisensor data integration. The two main steps in the data integration processing chain (shown in figure 1.1) are the registration of the images to the same reference frame (*image registration*),

and the fusion of the point values in the images that correspond to the same area on the ground (*data fusion*) (Pohl and van Genderen, 1998). Although only the geometric correction component of the data integration procedure is relevant to this research, a description of some data fusion techniques have been included in this chapter for completeness.

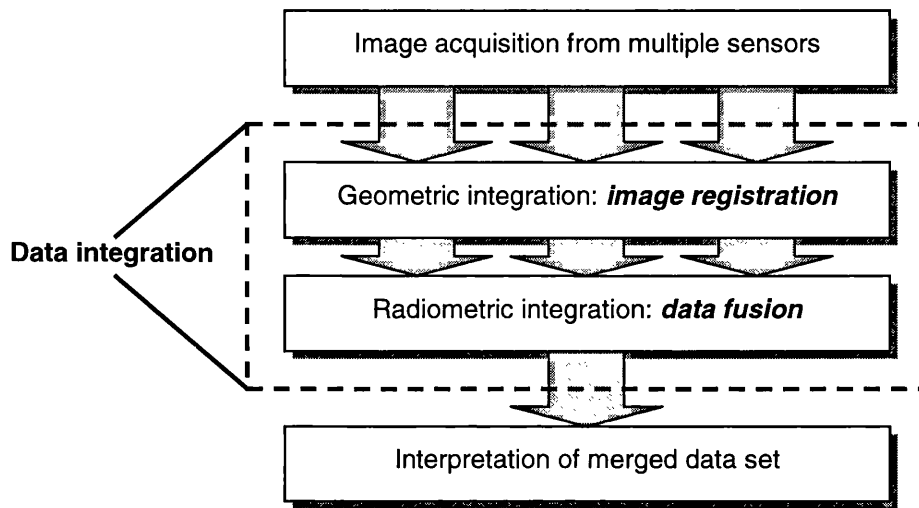


Figure 1.1 Data integration processing chain

Data integration can have a number of applications in fields as diverse as medical imaging, remote sensing, and robot vision. However, since research over the years has often been application oriented, many of the techniques developed have become application specific. This study is only concerned with the merging of remotely sensed images (images of the Earth acquired from airborne or spaceborne sensors) and therefore concentrates solely on research done in the field of Earth observation.

Remotely sensed Earth observation images have been available for a number of decades. Initially, images were acquired by cameras mounted on aircraft, but as the technology developed, new types of image acquisition systems were invented. Electro-optical devices, known as scanners, were developed, which are able to acquire optical images electronically without the need for photographic film – the images being stored digitally in a solid state memory device. Developments in radar technology led to radar imaging systems such as side looking airborne radar (SLAR) and synthetic aperture radar (SAR). In the early 1970s many of these different sensors were taken into low Earth orbit on board satellites and the first spaceborne remotely sensed images were acquired. This wide variety of different sensors means that images of the Earth can be acquired in many different wavelengths. This is extremely important since atmospheric absorption limits ^{the} range of wavelengths of electromagnetic radiation which can be used

to view the ground. Figure 1.2 shows a diagram of the electromagnetic spectrum and atmospheric absorption.

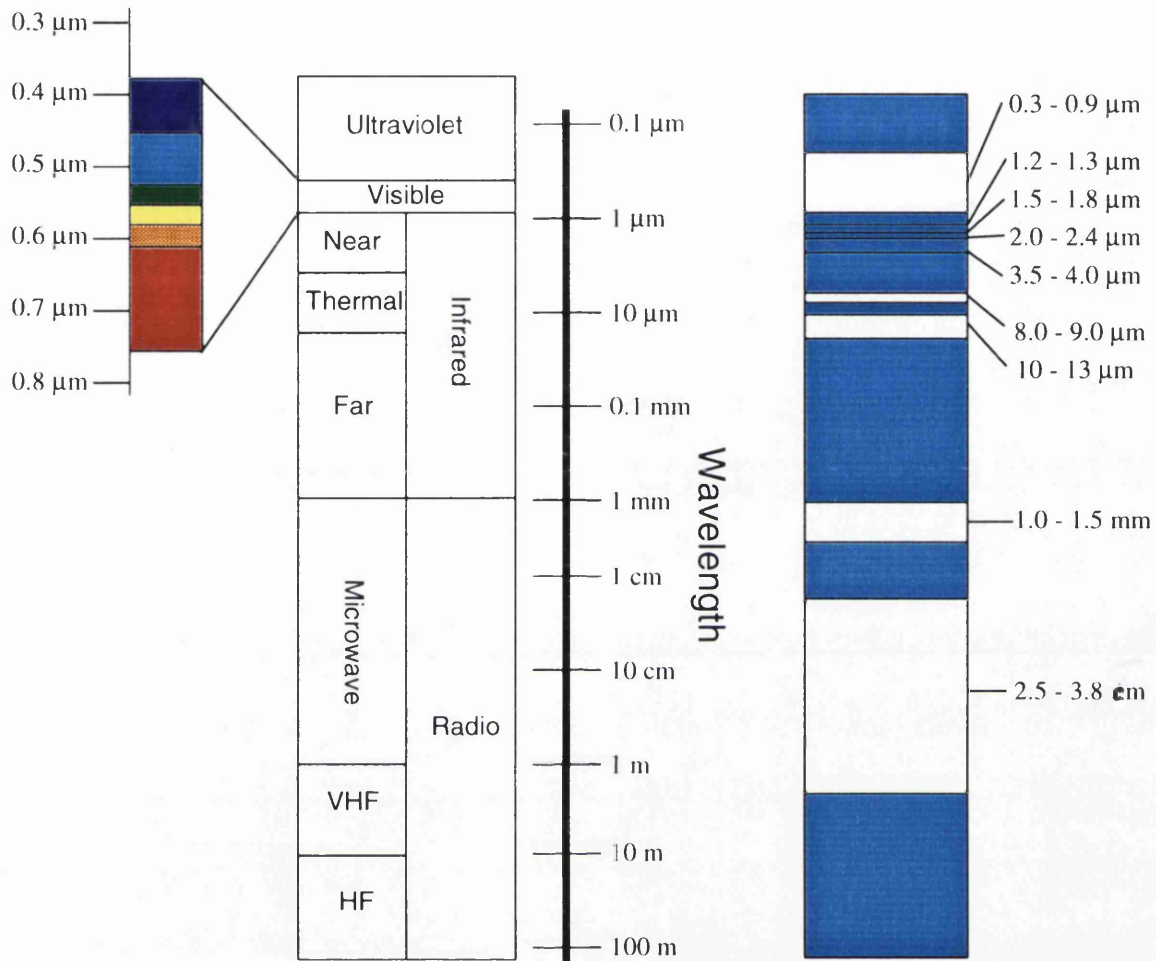


Figure 1.2 The electromagnetic spectrum and atmospheric absorption

The left hand side of figure 1.2 shows the electromagnetic spectrum split into different regions, and their corresponding wavelengths. The right hand side shows regions of high atmospheric absorption (in blue) and low atmospheric absorption, or "windows" (in white). Only radiation with a wavelength in the white regions is able to pass through the atmosphere to the sensor. It can be seen that Earth observation sensors are limited to acquiring data in only a few regions of the spectrum. Even so, this does give enormous opportunity for observing the Earth.

1.2 Image registration

Image registration is the process of geometrically transforming multiple images to the same reference frame. In the case where that frame of reference is a local geodetic

system (i.e. map projection), then the process is referred to as rectification¹. Figure 1.3 shows examples of image-to-image registration, and image-to-map rectification.

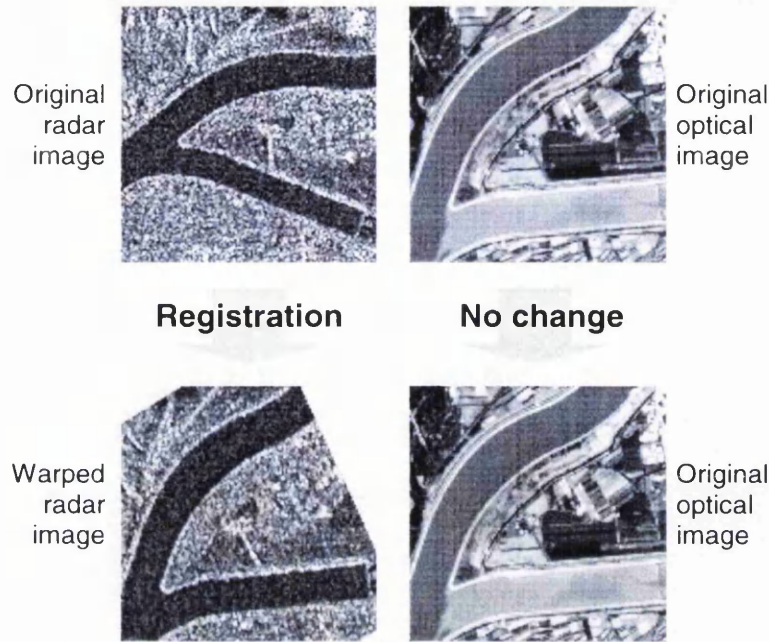


Figure 1.3 (a) Image registration

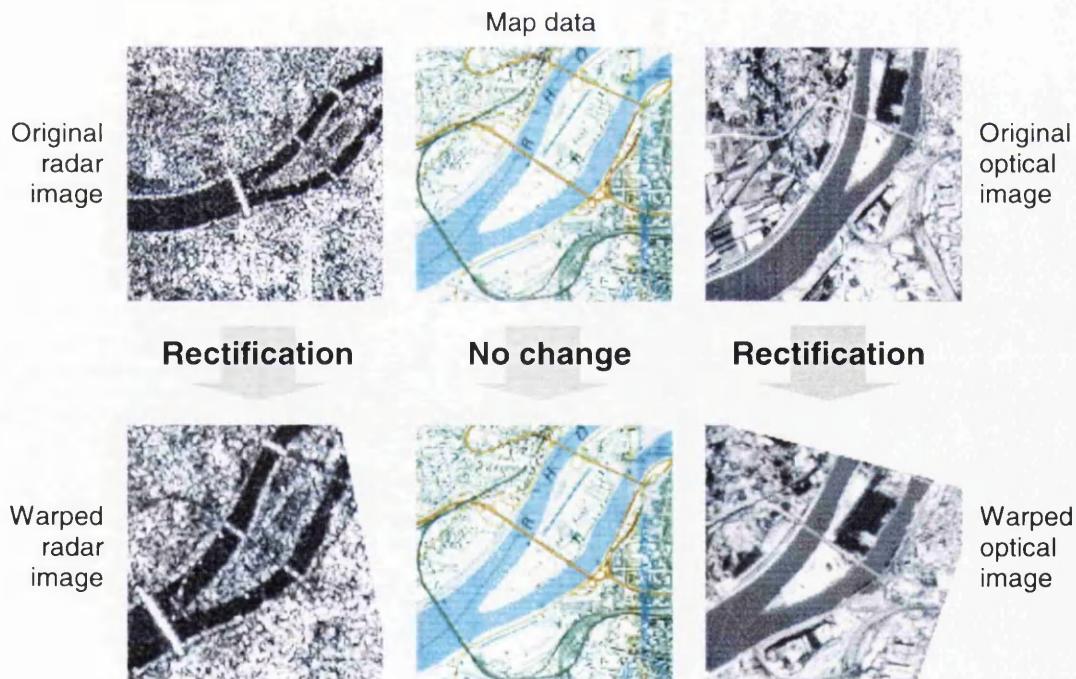


Figure 1.3 (b) Image rectification

Figure 1.3 (a) shows an example of image-to-image registration, where a SAR image is registered to a SPOT image. As a result, the SAR image has been displayed in the SPOT image internal co-ordinate system of rows and columns. Figure 1.3 (b) shows both a SAR image and a SPOT image being registered to a map. The pixels of the

¹ A glossary defining these and other terms is provided in § 1.8.

transformed images are expressed in ground co-ordinates (since this is the co-ordinate system of the map), so in this case the images have been rectified.

Although rectification removes the geometric effects of orientation from images, this is not the only type of distortion present in images. Other distortions are due to sensor anomalies, atmospheric refraction and terrain. All these different distortions seriously compromise the geometry of the image, and if accurate measurements are to be made from the image, then these distortions have to be eliminated. The process of removing terrain and sensor distortions is called orthorectification (Wolf, 1983), and is described in more detail in § 2.2.

Not all images have to be registered with each other in order to be merged; some groups of multiple images are acquired in a common frame of reference. The pixels making up these images can therefore be merged with each other without the need for image registration. These are images that are simultaneously acquired by the same sensor at the same time, the only difference between them being that they have been acquired in different parts of the electromagnetic spectrum. However, this is the exception to the rule, and in general images that have been acquired by the same sensor but at different times, or by different sensors, have to be registered before pixels can be merged. The accuracy of that registration will directly affect the quality of the data fusion, and ultimately the usefulness of the final data set.

1.3 Data fusion

Data fusion is the merging of point measurements from images that are in the same reference frame. It is not necessary that the data points being merged have been acquired by different sensors, only that they come from different images. There are numerous different ways of combining point measurements to highlight different aspects of the images in the final merged data set, but three commonly used techniques are (Pohl, 1996):

- ❑ arithmetic combination;
- ❑ statistical combination; and
- ❑ colour combination.

Arithmetic combination of data sets is simply the combination of pixel values from multiple input images using some function or formula to give a new output pixel value. Addition and multiplication operations can be used to enhance characteristics of the

data, whereas subtraction or division operations are useful for highlighting changes between multiple images. An example of data fusion, using arithmetic combination and images which originate from the same sensor, is the calculation of the normalized difference vegetation index (NDVI) from advanced very high resolution radiometer (AVHRR) images. AVHRR data is acquired in five channels, but channels one and two are commonly used to calculate vegetation indices. NDVI is a parameter sensitive to green vegetation, and hence a useful indicator of biomass. It is calculated using the formula:

$$NDVI = \frac{(CH_2 - CH_1)}{(CH_2 + CH_1)} \quad (1.1)$$

where CH_1 and CH_2 represent the reflectances of channels one and two respectively. Using these two quantities, the NDVI can be calculated for each pixel in the image, and displayed as a new image (figure 1.4). The new image is made up of data fused from the original input images.

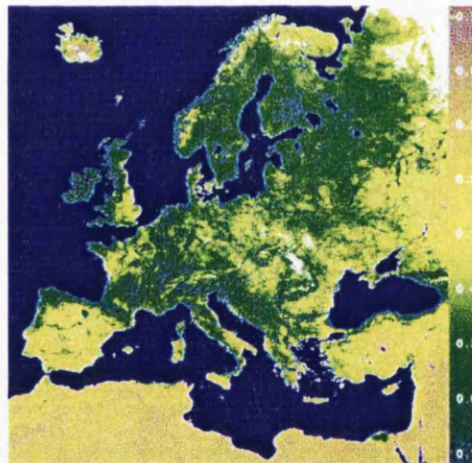


Figure 1.4 NDVI image of Europe for September 1996 (Gredel and Schröter, 1997)

An example of statistical combination of point measurements is principal component analysis (PCA). PCA is generally used to enhance multispectral data sets for improved interpretation or classification by removing some of the redundancy between the images from successive channels. In other words, PCA aims to reduce an n -channel data set to a data set containing less than n channels, or components. A detailed description of PCA is outside the scope of this thesis, but further details can be found in Lillesand and Kiefer (1994).

The last method of data fusion described here is colour combination, where three data sets are each assigned to the red, green and blue channels of a colour image, giving a

colour composite image. An example of colour combination using Shuttle Imaging Radar-C/X-band Synthetic Aperture Radar (SIR-C/X-SAR) images is shown in §1.4. A further method of combining point measurements using colour combination is the intensity-hue-saturation transform, where the different data sets are assigned to the intensity, hue and saturation components of a colour image. This is especially useful for image sharpening, an example of which is also given in §1.4.

One of the most important aspects of data fusion is that it relies heavily on the accuracy of the registration of the data sources being fused (Dai and Khorram, 1998). This is not an issue with images that come from different channels of the same sensor and have been acquired simultaneously (as with the AVHRR data above), but for multisensor images, or images from the same sensor but acquired at different times, it is essential that the images are accurately registered to each other in order to produce an accurate result. This question of accuracy can be exemplified using SPOT XS (multispectral) images that have been deliberately misaligned (see figure 1.5).

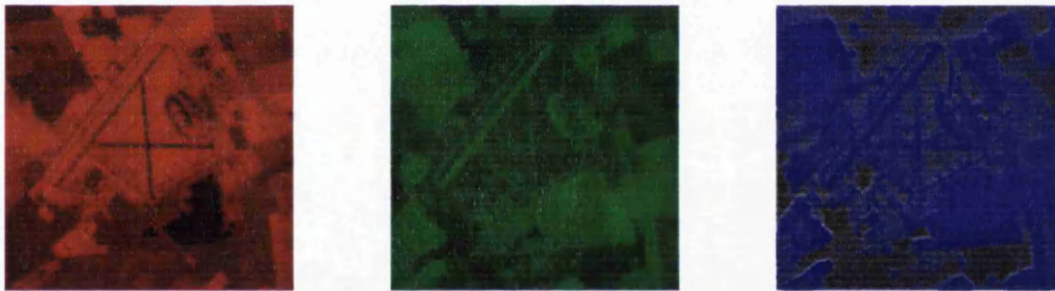


Figure 1.5 (a) Three channels of SPOT multispectral data

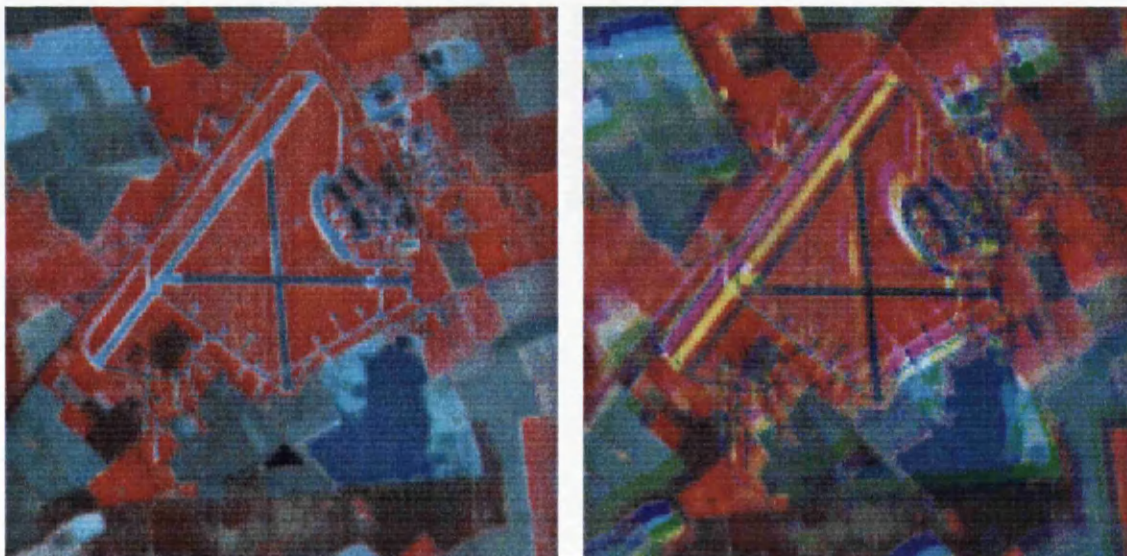


Figure 1.5 (b) Correctly aligned SPOT XS channels (left), misaligned channels (right)

Figure 1.5 (a) shows three SPOT images of an airfield in Central England corresponding to channels 1, 2 and 3. In figure 1.5 (b) the image on the left shows the combined image

with the channels correctly aligned, but the image on the right has been misaligned by three pixels. It can clearly be seen that although the misalignment is quite small, it makes a large difference to the interpretability of the image.

1.4 Applications of multi-image data sets

The product of data integration, the multi-image data set, can have many practical applications. It is these applications that make data integration a useful research tool, rather than just an interesting academic exercise. This section describes some important practical uses of data integration, such as image enhancement, change detection and improved vegetation mapping, which illustrate how multi-image data sets can be applied to "real world" problems.

Figure 1.6 (a) shows two SIR-C/X-SAR images of a region of the Brazilian rainforest near the town of Manaus. The image on the left was acquired on 12 April 1994 and the image on the right was acquired on 3 October 1994. The images show the Rio Solimoes (the main dark region) and the surrounding rainforest (the lighter regions). There are obvious differences in brightness between the images, which reflect changes in the scattering of the radar pulses. These brightness differences are indicative of flooding: a flooded forest has a higher backscatter than an unflooded region. The extent of the flooding is much greater in the April image than in the October image, and corresponds to the annual 10-metre rise and fall of the Amazon River.

Figure 1.6 (b) shows the change in the April and October images and was created by determining which areas had significant decreases in the intensity of radar backscatter. These areas, which appear blue, show the increase in the extent of flooded forest, as the level of the Amazon River rises. Using these radar images in conjunction with ground surveys allows the creation of an inundation map for this region of the Amazon (figure 1.6 (c)).

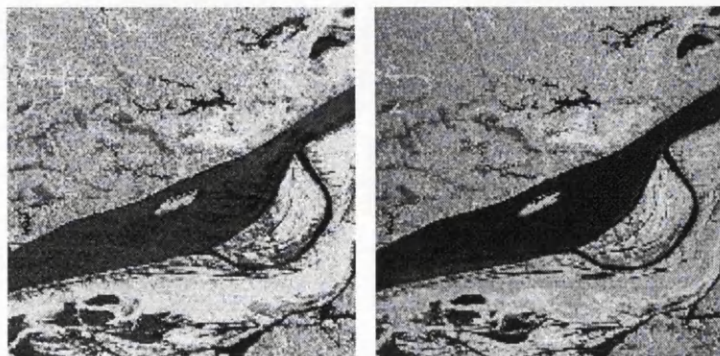


Figure 1.6 (a) SIR-C/X-SAR images of Manaus, 12/04/94 & 3/10/94 respectively (© NASA)



Figure 1.6 (b) Difference image representing changes in flooding (© NASA)

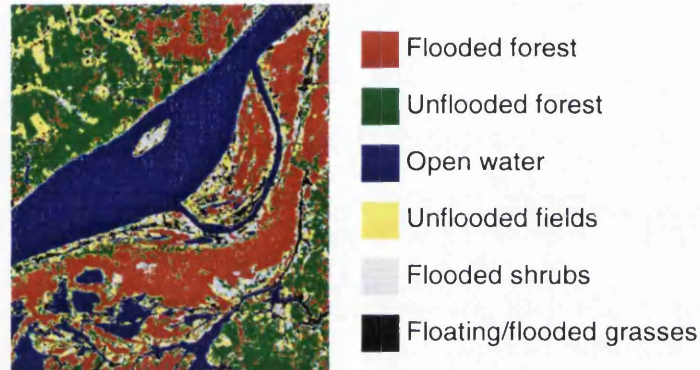


Figure 1.6 (c) Inundation map of Manaus, Brazil (© NASA)

A second application of multi-image data sets once again uses data from the SIR-C/X-SAR instrument acquired over the Amazon basin. Figure 1.7 shows three images of the Rodonia region in western Brazil, acquired at the same time by the same instrument, but at different frequencies and polarizations². Figure 1.7 (a) shows an image acquired in the L-band with HV polarization, figure 1.7 (b) shows an image acquired in the C-band with HV polarization, and figure 1.7 (c) shows an image acquired in the X-band with VV polarization.

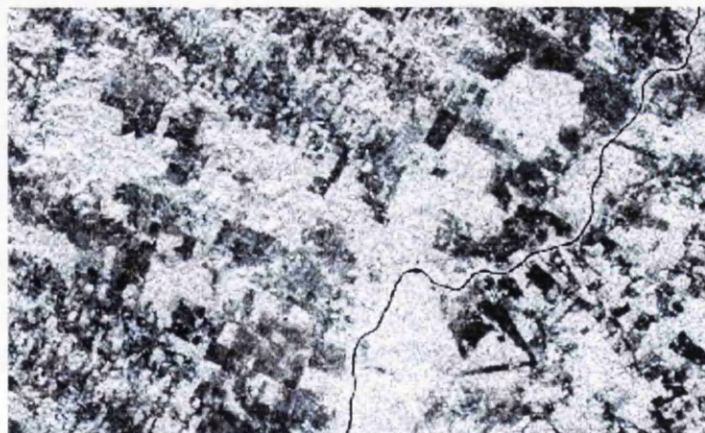


Figure 1.7 (a) L-band, HV polarization (© NASA)

² Radar bands and polarizations are described in chapter 3.

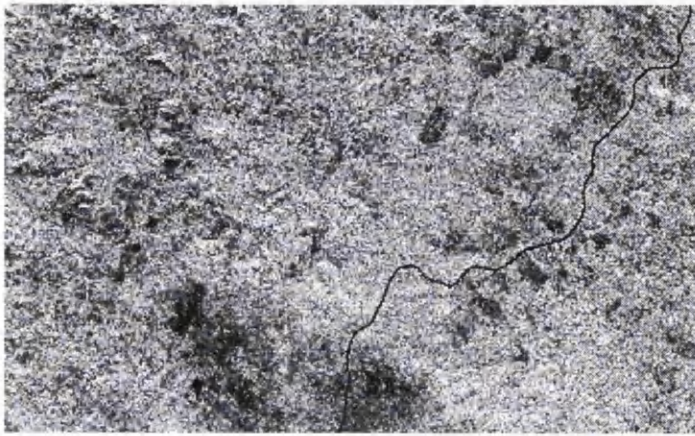


Figure 1.7 (b) C-band, HV polarization (© NASA)

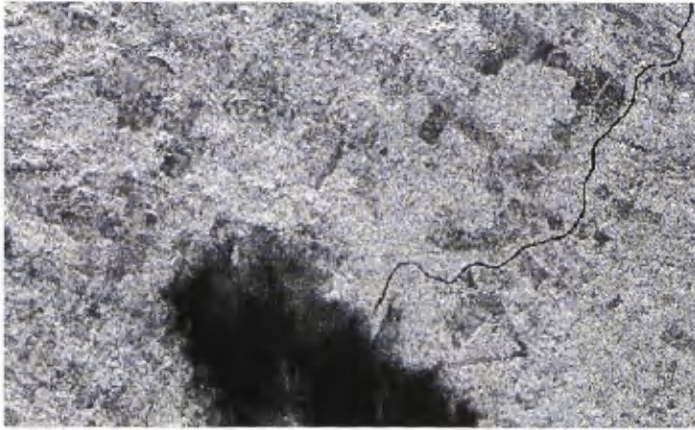


Figure 1.7 (c) X-band, VV polarization (© NASA)

The different frequencies and polarizations make it possible to observe different types of features on the ground. The L-band image clearly shows vegetation patterns and it is easily possible to differentiate clear cut areas from virgin rainforest. The X-band image contains a dark feature which is an intense thunderstorm. Usually radar is able to penetrate cloud cover, but shorter wavelengths have more difficulty doing so. In this example the X-band radiation is actually being scattered by the raindrops within the storm, and not the cloud itself. The amount of scattering is dependent on the size of the drops and the intensity of the rainfall, which means that X-band images can be used for the estimation of rainfall rates.

These images can be combined into a single image by fusing the data into a red-green-blue (RGB) false colour composite (figure 1.8). The X-band image is blue in the colour image, the C-band image is green, and the L-band image is red. This multi-frequency radar image shows rapidly changing land use patterns and also demonstrates the capability of the different radar frequencies to detect and penetrate heavy rainstorms. When combined in the colour image, the rain storm appears red and yellow. The pink areas are virgin rainforest, and the blue and green patches are areas where the forest has been cleared for agriculture.

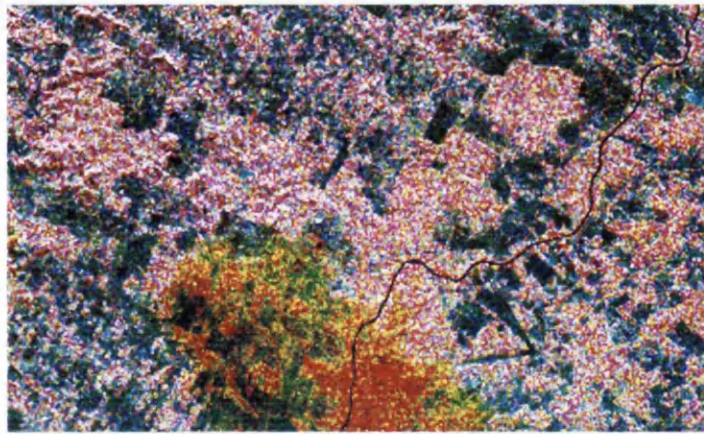


Figure 1.8 False colour composite of Rodonia, western Brazil (© NASA)

A third application of data integration is the use of high resolution images to sharpen lower resolution colour (multispectral) images using the intensity-hue-saturation (IHS) transformation. A sharpened colour image is generated by assigning the intensity, hue and saturation to the data from three sources: the data from the higher resolution sensor replaces the intensity channel of the lower resolution image, whilst the hue and saturation channels of the lower resolution image remain unchanged. Figure 1.9 (a) shows two Indian Remote Sensing Satellite (IRS-1C) images of Frankfurt airport: the one on the left is 5.8m resolution panchromatic and the one on the right is 23m resolution multispectral. The first stage of the data integration process is image registration, and figure 1.9 (b) shows the multispectral image registered to the panchromatic image. In figure 1.9 (c) the images have been fused to produce an artificially sharpened multispectral image. The intensity layer of the multispectral image has been replaced with the higher resolution panchromatic image.

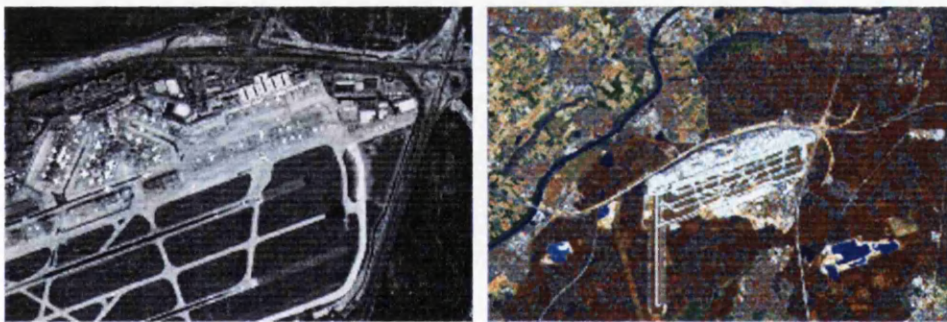


Figure 1.9 (a) 5.8m panchromatic and 23m multispectral IRS-1C images (© Euromap)

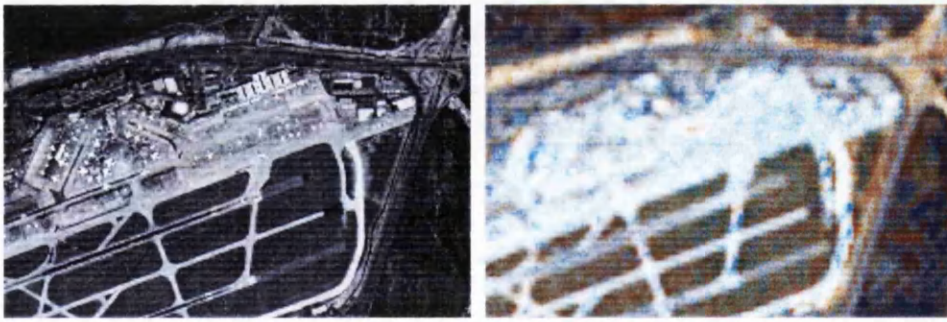


Figure 1.9 (b) 5.8m panchromatic image with registered 23m multispectral image

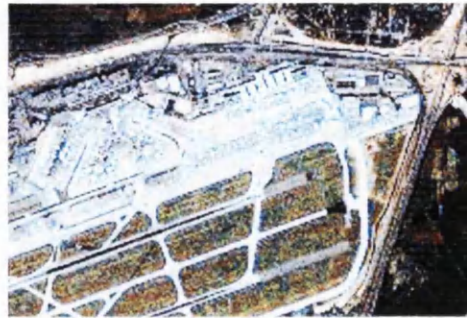


Figure 1.9 (c) IHS merged image

1.5 Problems associated with data integration

Two major requirements associated with any use of Earth observation data are accuracy and speed of delivery. For any image product to be of real use to the end user, it must be accurate and it must be available when it is needed. These two requirements are especially relevant to multi-image remotely sensed data products. Integrating different images can slow down the delivery of the product, and if the integration is not done well, the final product may not be accurate enough for its intended purpose. This can be illustrated by two examples: crop monitoring and disaster management.

Crop monitoring is important the world over. In Europe, Earth observation data from satellites and aircraft is used to monitor the production of various crops from individual countries in order to combat agricultural fraud. In other parts of the world, remotely sensed images are used to assess the health of crops to improve yield. In both cases, a greater diversity of crops can be monitored with more accuracy if the target area is observed with more than one type of sensor (Schistad-Solberg et al., 1994), but this will only be true if the images are accurately registered (see figure 1.5 above). Inaccurate registration will lead to a multi-image product that would possibly be worse for observing crops than using a single image in isolation. With manual image registration (described in chapter 2) the accuracy of the final image product is very dependent on the time the user spends performing the registration.

More and more nowadays, remotely sensed images are being used for disaster management (Buongiorno et al., 1997; Barbieri et al., 1997). Satellite images acquired before and after an earthquake for example, can help the rescue services on the ground target their efforts to where they are most needed. The same is also true for many other types of disasters, both manmade and natural: remotely sensed data can be used to measure the location and extent of oil slicks, floods, forest fires etc. The value of the data product will always be higher if more than one data set has been combined, since either more features will be visible on the ground, (see figure 1.8, above), or the images will accurately show changes that have taken place on the ground (see figure 1.6, above). However, in all cases of disaster monitoring, the value of the image product will decrease very quickly with time. Manual image registration is a very user-intensive and time consuming process which requires knowledge, skill and experience. It will slow down the delivery of the multi-image data product considerably.

The above two examples illustrate that manual image registration is an insufficient tool when data integration is applied to "real world" problems. Thus the challenge for the research community is to develop tools for integrating data sets quickly and accurately, which therefore means *automatically*.

1.6 Automatic multi-image data integration

The two reasons cited above supporting the concept of automatic image registration are the need for high accuracy and prompt delivery to the user. A further reason for researching automatic image registration concerns the sheer volume of data acquired by Earth observation sensors. The Committee for Earth Observation Satellites has identified 126 current missions with 217 instruments existing or planned. By the 21st century it is anticipated that 20 terabits (20×10^{12} bits) of data will be acquired every day (the equivalent of 600 000 bibles being produced every day) (Harris, 1997). The best way to exploit all this data to its full potential is to combine it into integrated multisensor data sets, but it will not be physically possible to do this without automatic processing techniques.

The nature of the automation will depend on the stage in the processing at which the integration actually takes place. There are three stages in the processing chain where data integration can take place:

1. the image acquisition stage, by specialized software onboard the aircraft/satellite (Prechtel and Bringmann, 1998);

2. the pre-distribution stage, by trained operators; and
3. the post-distribution stage, by the end user.

If the integration of data sets takes place at the image acquisition stage, then the processing must be fully automatic and very robust since, due to logistics, there is no opportunity for human operators to intervene. Integration at the pre-distribution stage, however, requires far less robust procedures since trained operators are able to monitor the processing and correct any blunders made. In addition, the processing does not have to be completely automatic, since operators are available to perform some of the more intricate tasks that are difficult to automate. Integration at the end user stage falls somewhere between the two previous categories - a human operator may be available to correct blunders, but that operator will probably only be able to perform simple tasks. Therefore, when developing automatic processing algorithms it is important to define the level of automation required of those algorithms based on how the algorithms will be used.

On a more practical level, the degree of automation often depends on which steps in the processing chain can be automated without a loss of quality in the result. If the automatic processing chain produces a less accurate result, or a result that takes too long to achieve, then it may be better to have a manual step in that processing chain. Consequently, the most efficient processing (in terms of time and accuracy) may be a semi-automatic one, where the user is available to make some decisions, but the computer performs most of the complex and laborious tasks.

1.7 Discussion

This chapter has introduced the subject of data integration and illustrated its importance with some possible applications. The two main steps in data integration, image registration and data fusion, are each very active and important research topics in their own right. To research both of these topics in one piece of work would be far too much to take on, so the remainder of this thesis is concerned only with image registration: the geometrical and spatial issues associated with data integration. As the next chapter will show, image registration is a mature subject which has benefited greatly from the vast amount of research done over the years. However, there are still problems associated with image registration that have not yet been fully resolved, a particularly troublesome one being the automatic registration of multi-sensor data sets. Automatic registration of two similar images (such as a pair of aerial photographs) is a solved problem (Day and

Muller, 1989; Heipke, 1997), but a robust method of automatically registering pairs of dissimilar images is far from available.

This thesis sets out to investigate the possibilities of automatically registering data from two very different spaceborne Earth observation sensors as accurately as possible, but with a minimum of human intervention. The image types that were chosen to be used in this study are SAR data from the sensor onboard the ERS-1 platform, and panchromatic SPOT images from the high resolution visible (HRV) sensor onboard the SPOT platform. The reason for choosing these two image types in particular is due to their very similar ground pixel size (12.5m for SAR and 10m for SPOT) and their very different methods of imaging the Earth's surface (described in more detail in chapter 3), which leads to some very interesting differences in radiometry and geometry between the two image types. A detailed description of the work presented in this thesis is given at the end of the next chapter, after the subject of image registration has been described in more detail.

1.8 Glossary of terms

Table 1.1 gives the definitions of a number of terms used in this thesis. There is some discussion about the exact definitions of some of the terms, so in this study the terms are used according to how they are defined in this table. Cross references to other definitions are indicated with *italics*.

Term	Definition
Accuracy	Accuracy is the relationship of a set of features to a defined reference system and is expressed as the <i>rms error</i> of a set of derived points.
Blunder	See <i>Error</i>
Check Point	A well-defined <i>ground reference</i> point used for checking the <i>accuracy</i> of a <i>geometrically corrected</i> image or image mosaic. Check points must not be the same as <i>GCPs</i> . Adapted from Wolf (1983).
Digital Elevation Model (DEM)	A digital, raster representation of land surface elevation above sea level. DEM is used in preference to digital terrain model (DTM) because the term 'terrain' implies attributes of the landscape other than elevation. Adapted from Burrough (1986)

Discrepancy	A discrepancy is the linear distance between a point on the image and a <i>check point</i> . A discrepancy is not the same as a <i>residual</i> , because a discrepancy is an <i>error</i> at each point measured using a reference point known to a higher order of accuracy.
Error	<p>Three classes of error are commonly recognised:</p> <p>A <i>random error</i> is not predictable at any given location but the population of random geometric errors commonly follows a normal (Gaussian) probability distribution. If random errors are normally distributed the mean error is zero for a large sample of points.</p> <p>A <i>systematic error</i> is predictable at any given location once it has been identified and its pattern of variation is understood. For a large sample of points, a mean error that is not zero usually indicates the presence of a systematic error.</p> <p>A <i>blunder</i> is a (large) error at one location arising from an observer error or equipment fault whilst marking the location or recording its co-ordinates.</p>
Geocoding	Synonym for <i>orthorectification</i> , but more commonly used when discussing SAR data.
Geometric correction	Informal term for <i>rectification</i> or <i>orthorectification</i> .
Ground control point (GCP)	A well-defined point used for orientation and <i>rectification</i> . The position of a GCP is known both in <i>ground reference</i> co-ordinates and in the co-ordinates of the image to be corrected. If 2D (x,y) ground reference co-ordinates are given, it is a horizontal or planimetric GCP; if the height (z co-ordinate) is known, the point is a vertical GCP.
Ground Reference	The source used to obtain ground reference co-ordinates for a <i>ground control point</i> or <i>check point</i> . May be a topographic map, a field survey by triangulation, a geodetic bench mark, a field survey by GPS, or a <i>geocoded</i> image. The ground reference will provide co-ordinates in the national map projection.

Interpolation	Method used to estimate a pixel value for a corrected image grid, when resampling from pixel values in the original grid. Common methods are nearest neighbour, bilinear interpolation and cubic convolution.
Orientation	<p>Orientation can have three components.</p> <p>Interior orientation establishes precise relationships between a real image and the focal plane of a perfect imaging system.</p> <p>Relative orientation establishes precise relationships between the focal planes of a perfect stereopair to establish a precise stereomodel.</p> <p>Exterior orientation establishes precise relationships between the focal plane co-ordinates of an image (or stereomodel) and a geographic co-ordinate system (map projection).</p>
Orthorectification (orthocorrection)	<i>Rectification</i> of an image (or image stereo pair) using 3D <i>ground reference</i> and a <i>DEM</i> to position all image features in their true orthographic locations. The process eliminates displacements due to image geometry and topographic relief, and results in an image having the same geometric properties as a map projection (Wolf, 1983).
Pixel size	Distance represented by each pixel in an image or <i>DEM</i> in <i>x</i> and <i>y</i> components. Pixel size can be expressed as a distance on the ground or a distance on scanned hardcopy (e.g. microns). It is not a measure of <i>resolution</i> .
Polynomial rectification	<i>Rectification</i> of an image to a <i>ground reference</i> using horizontal <i>ground control points</i> . It assumes that the local distortion of the image is uniform and continuous since it ignores effects of terrain.

Precision	<p>The precision of a <i>GCP</i> or <i>check point</i> is the standard deviation of its position (in <i>x</i>, <i>y</i> and <i>z</i>) as determined from repeated trials under identical conditions.</p> <p>Precision indicates the internal consistency of a set of data and is expressed as the <i>standard deviation</i> (σ).</p> <p>Note: Data can be precise yet inaccurate; precision is not used when comparing a set of data to an external reference, <i>RMSE</i> is used to express this.</p>
Rectification	The process of resampling pixels of an image into a new grid which is referenced to a specific geographic projection, using a spatial transformation (matrix). The resampling is achieved through <i>interpolation</i> .
Registration	<i>Rectification</i> of an image to conform to another image.
Residual	<p>A residual is the linear distance between a fixed reference point [ground control point] and the position determined by the transformation applied to the observed data to give a best fit to the reference points.</p> <p>Note: This is not the same as a <i>discrepancy</i> because the computed error of a residual is based only on the internal (statistical) consistency of a set of points and not on comparison to independent locations known to higher accuracy.</p>
Resolution (resolving power)	The smallest visible separation between similar objects that can be clearly reproduced by a remote sensing system – usually expressed as the maximum number of line pairs per unit length (Light, 1993).
Root mean square error (RMSE)	<p>The square root of the average of the squared <i>discrepancies</i> or <i>residuals</i>: $\sqrt{\frac{1}{n} \sum_{i=1}^n d_i^2}$ where <i>d</i> is the measured discrepancy or residual in <i>x</i>, <i>y</i> or <i>z</i>. Adapted from ASPRS (1988).</p>
RMSE (Absolute)	<i>RMSE</i> based on <i>check points</i> obtained from a <i>ground reference</i> of recognised higher <i>accuracy</i> (EC, 1997).

RMSE (Relative)	<i>RMSE</i> based on <i>check points</i> extracted from another <i>geocoded</i> image. In practice the <i>RMSE</i> of the GCP <i>residuals</i> is also used as a measure of relative error (EC, 1997).
Standard Deviation (σ)	The square root of the variance of n observations, where the variance is the average of the squared deviations about the estimate of the true mean value.
Tie points	Points that appear on the overlap area of adjacent images. They are used for <i>orientation</i> and aerotriangulation.
Tolerance	The tolerance is the permissible degree of <i>error</i> in a geometrically corrected image or mosaic as determined using a well distributed set of <i>check points</i> . Tolerance is specified with two values: a) the maximum allowable <i>RMS error</i> of all check points b) the maximum allowable <i>discrepancy</i> at any check point.

Table 1.1 Definitions of some terms used in this thesis

Chapter 2 IMAGE REGISTRATION

2.1 Introduction

Chapter 1 introduced the subject of data integration and explained why it is a two stage procedure (image registration and data fusion). The aim of this chapter is to explain the subject of image registration in more detail. However, in order to get a better understanding of image registration, this chapter begins with a description of rigorous geometric correction of remotely sensed images. It then goes on to describe the process of manual image registration, followed by automatic image registration. After the introduction and description of these concepts, a detailed outline of the proposal of this thesis is presented.

2.2 Rigorous geometric correction of images

2.2.1 Introduction

This section describes the orthorectification of remotely sensed images. The first part describes the photogrammetric techniques used to correct optical images (those images acquired in the visible or infrared part of the electromagnetic spectrum) whilst the second part describes the correction of radar images. Historically, the accurate geometric correction of radar images is referred to as geocoding rather than orthorectification, but in general these two terms are synonymous.

2.2.2 Orthorectification of optical images

To orthorectify optical images, it is essential to know:

- ❑ the geometry of the imaging system;
- ❑ the position and attitude of the imaging system in relation to the ground; and
- ❑ the shape of the ground.

Sensor geometry

The geometry of the imaging system is defined by the type of sensor used. In optical remote sensing there are generally three groups of imaging systems: a frame camera, a

pushbroom scanner and a whisk-broom scanner. Frame cameras, used in aerial photography, acquire the whole of the image in a single exposure (there is no time element involved in the acquisition) and the image acquisition can be modelled by a central projection. A pushbroom scanner acquires one line of the image at a time, and builds the image by combining numerous lines. Each line is time independent, but a time factor must be considered between subsequent lines. An example of a pushbroom scanner is the HRV sensor on the SPOT satellite. A whisk-broom scanner records each pixel in the image at a different time, and therefore is the most difficult to model mathematically. An example of a whisk-broom scanner is the Thematic Mapper (TM) instrument of the Landsat series of satellites. The frame camera will be used throughout the rest of this section to illustrate the orthorectification processing chain since it is the simplest to describe, but the same procedures can be applied to the other types of sensors, albeit with more difficulty.

Platform position and attitude

The position and attitude of the imaging system in relation to the ground can be determined by either using GCPs, or data from a combined global positioning system (GPS) inertial navigation system (INS) onboard the platform. The traditional method of determining position and attitude using GCPs is more precise than using GPS/INS since the latter is still in the research stage, and at present limited by technology. By simply locating points in the image which represent known points on the ground and using knowledge of the sensor geometry, it is possible to determine a precise location for the sensor in space.

Shape of the ground

The shape of the ground is defined by a DEM, a representation of the ground that uses a grid of points. A DEM can be determined from a number of different methods, but the most commonly used methods are stereo photography or ground survey.

Using these three pieces of information (sensor details, position of sensor and shape of terrain) it is possible to correct for distortions in an aerial photograph caused by viewing angle, inconsistencies in the shape of the lens, and relief. The lens distortions are corrected using information from a camera calibration certificate, which gives details of inconsistencies in the camera lens. The cause for the distortions due to relief is shown in figure 2.1.

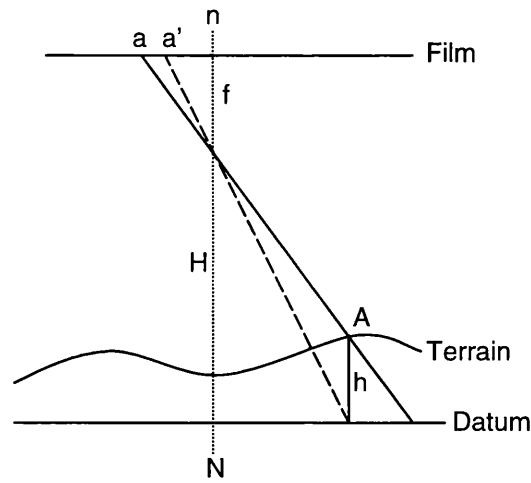


Figure 2.1 Distortion due to relief in aerial photographs

In figure 2.1, the point A on the ground is represented by point a on the image. However, the true orthographic location of A should be represented by a' , hence there is a radial displacement of A in the image from a' to a by a distance dr , where dr is given by:

$$dr = \frac{r \cdot h}{H} \quad (2.1)$$

where r is the radial distance of a from the principal point, H is the altitude of the sensor and h is the height of the terrain at A. For an image acquired by a camera at an altitude of 10km, and a point on the ground 500m above the datum and 4km from the nadir, dr (in ground units) is 200m, which is obviously a significant amount that has to be corrected. Figure 2.2 shows how these displacements can be removed.

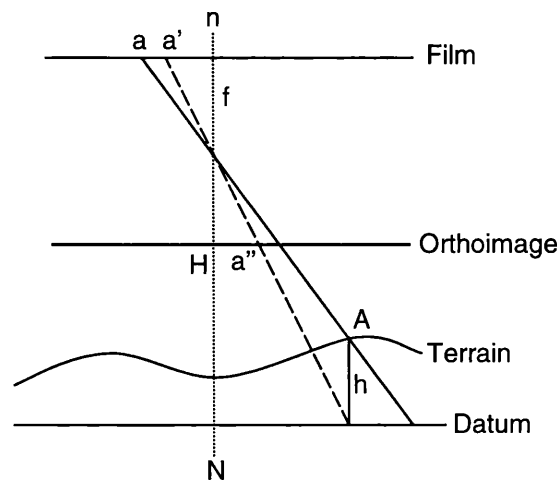


Figure 2.2 Orthoimage production

Using the DEM, the height for every point on the ground, A_i , is known, so the displacement due to terrain of each point a_i on the image can be calculated. Therefore,

for all A_i , the location of a_i' , the true orthographic location of A_i , can be determined. When a_i' is reprojected onto the orthoimage, the points A_i are represented by a_i'' . Thus the orthoimage shows the true orthographic locations of all the points in the original image.

The same ^{procedure} can be applied to the correction of satellite images as well as aerial photographs. However, for satellite data, the precision and accuracy of the DEM and the GCPs does not have to be quite as high, since (up to now) the resolution of the satellite sensor is much lower than an aerial camera.

2.2.3 Geocoding of SAR images

The term geocoding is synonymous with orthorectification, although where orthorectification is generally used to refer to the correction of optical images, geocoding is used with radar images. Therefore, the term geocoding describes the process of removing terrain induced errors from radar images, and resampling those images to a geodetic frame of reference. This procedure has been described by a number of researchers, but a clear and concise summary is given by Curlander et al. (1987).

To geocode any SAR image, the supplementary information required is a DEM, the state vectors (position and velocity vectors) of the sensor at the time of image acquisition, and a couple of GCPs. As with orthorectification, the final geometric quality of the geocoded image will depend directly on the precision and accuracy of this supplementary information. If the state vectors are of a high enough quality (as is the case with ERS-1/2) no GCPs are required for the processing, unlike orthorectification of optical images where GCPs are essential in determining the location of the sensor. For other SAR sensors, such as Radarsat or JERS-1, one or two GCPs are required – still far fewer than for orthorectification of optical data. Using this information, it is possible to determine the absolute position (x, y, z in global co-ordinates) of all of the pixel values in the corresponding SAR image. This is done by solving three simultaneous equations (Curlander et al., 1987):

1. The Earth model equation,³

$$\frac{x^2 + y^2}{R_e^2} + \frac{z^2}{R_p^2} = 1 \quad (2.2)$$

where (x, y, z) is the location of the target on the ground, R_e is the mean equatorial Earth radius and $R_p = (1 - 1/f) R_e$ where f is a flattening factor.

2. The SAR Doppler equation,

$$f_D = \frac{2}{\lambda R} (\bar{V}_s - \bar{V}_t) \cdot (\bar{R}_s - \bar{R}_t) \quad (2.3)$$

where f_D is the Doppler parameter, λ is the radar wavelength, R is the sensor to target slant range, \bar{V}_s and \bar{V}_t are the sensor and target velocity vectors, and \bar{R}_s and \bar{R}_t are the sensor and target position vectors.

3. The SAR range equation,

$$R_{ij} = \sqrt{(\bar{R}_s - \bar{R}_t) \cdot (\bar{R}_s - \bar{R}_t)} \quad (2.4)$$

where R_{ij} is the slant range from sensor to target for pixel (i, j) .

The fact that GCPs are not usually needed for ERS images means that the geocoding of these images is essentially an automatic procedure.

2.2.4 Conclusions

This section has given a brief overview of the fundamental principles of rigorous geometric correction of aerial and satellite remotely sensed images. It is important to note that the methods presented require either very accurate supplementary data, or human intervention, or both. The next section now goes on to describe the simplest method of registering a pair of images: manual image registration.

2.3 Two dimensional image registration

2.3.1 Introduction

Image registration (introduced in section 1.2) is a well defined procedure, and an example is shown in figure 1.3 (a). The processing chain of manual image registration is shown in figure 2.3. The next four sections of this chapter describe each of these four steps in more detail.

³ The Earth model equation represents an oblate ellipsoid, which is modified during the processing using the DEM to generate a more realistic model of the Earth's surface.

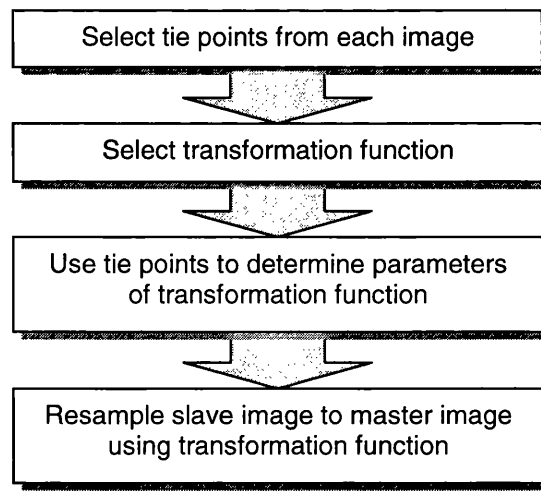


Figure 2.3 Manual image registration processing chain

2.3.2 Tie point selection

The first step in the processing chain is to manually select tie points from each image. A tie point is a point that can be explicitly and accurately identified in both images, and therefore used to ‘tie’ the two images together at that point. The minimum number of tie points required depends upon the type of transformation function which is being used to register the images.

2.3.3 Transformation functions

Selection of the transformation function is the second step in the processing chain, and depends on the types of distortions present in the images that are being registered. For Earth observation images, the transformation function is a two dimensional polynomial; the simplest is the similarity transformation, which is a polynomial of order one and with only four unknowns. The similarity transformation will only correct for differences in rotation, scale (which must be equal in both dimensions) and translation between the two images. Therefore, in the transformation process, angles and parallel lines are preserved. The transformation can be expressed as two simultaneous equations:

$$X = ax + by + c \quad (2.5a)$$

$$Y = bx - ay + d \quad (2.5b)$$

where x and y represent the co-ordinate system for image 1, X and Y the co-ordinate system for image 2, and a , b , c , d are the parameters of the transformation function. Solution of these equations for a , b , c , d requires two values of each of X , Y , x , y , which is the same as two tie points, (X_1, Y_1) and (x_1, y_1) , and (X_2, Y_2) and (x_2, y_2) .

The next simplest transformation, and the one most used in registering Earth observation images is the general first order polynomial transformation, the affine transformation. This transformation corrects for differences in rotation, scale (which may be different in each dimension) and rotation. Parallelism is preserved in the affine transformation, but angles are not. The two simultaneous equations for the affine transformation are:

$$X = ax + by + c \quad (2.6a)$$

$$Y = dx + ey + f \quad (2.6b)$$

To solve for the six unknowns (a to f) a minimum of three tie points are required. Another polynomial transformation of relevance to Earth observation image registration is the general second order polynomial transformation. Since with this transformation neither parallelism nor angles are preserved, it is sometimes known as rubber-sheeting. The two simultaneous equations for the second order transformation are:

$$X = a + bx + cy + dxy + ey^2 + fx^2 \quad (2.7a)$$

$$Y = g + hx + iy + jxy + ky^2 + lx^2 \quad (2.7b)$$

These equations contain twelve unknowns, and therefore require a minimum of six tie points to solve them. The minimum number of tie points, n_{\min} , required to solve any polynomial transformation function of order t is given by the equation:

$$n_{\min} = \frac{(t+1)(t+2)}{2} \quad (2.8)$$

However, using only the minimum number of tie points is unwise, since it allows no room for error. If one of the tie points is measured wrongly (known as a blunder) then there is no way of knowing this. If more than the minimum number of tie points are used to solve the equations (using a least squares method) then any blunders will be obvious in the result. Although there are no hard and fast rules, it is good practice to select as many tie points as possible, and make sure they are well distributed across the images.

To help decide which is the best transformation to use in a particular situation, or which combination of tie points will give the best result, it is necessary to have some indicator of registration accuracy. The indicator used in image registration is known as the

residual. A residual is defined as being the distance between where a transformed point would be expected to be found, and where it is actually found. Very often, the residuals of the tie points are used to assess the quality of the transformation (so long as more than the minimum required number of tie points are used in the calculation), but as will be shown below, this is not reliable test - instead the residuals of check points must be calculated. A check point has two properties: firstly, it is a point which can be accurately identified in both the images that are being registered, and secondly, it is not used in the calculation of the parameters of the transformation. As with the selection of tie points, it is good practice to select a large and well distributed set of check points. If the images have been registered well (i.e. a good selection of tie points have been used to generate the parameters of the correct transformation function) , then the residuals will be low. However, if the tie points were not selected well, or the wrong transformation function was used, then the residuals will be much higher. In addition to using the magnitudes of the residuals to quantify the quality of the registration, the vector component of the residuals can be used to detect systematic errors in the transformation. This is done by plotting the residuals and observing the distribution of their directions. In general, the distribution of directions would be expected to be fairly random, but if this is not the case (i.e. many of the residuals are aligned in a particular direction) then there may be a systematic error in the transformation. A systematic error could be due to the incorrect use of a particular transformation function, or a blunder in the processing, neither of which would necessarily be apparent by observing the magnitude of the residuals alone.

The importance of using check point residuals rather than tie point residuals when assessing the accuracy of a transformation can be illustrated with an example of image mosaicing using manual image registration. Figure 2.4 shows two images of Lake Kariba on the Zimbabwe-Zambia border acquired by a handheld camera onboard the Space Shuttle in 1991.

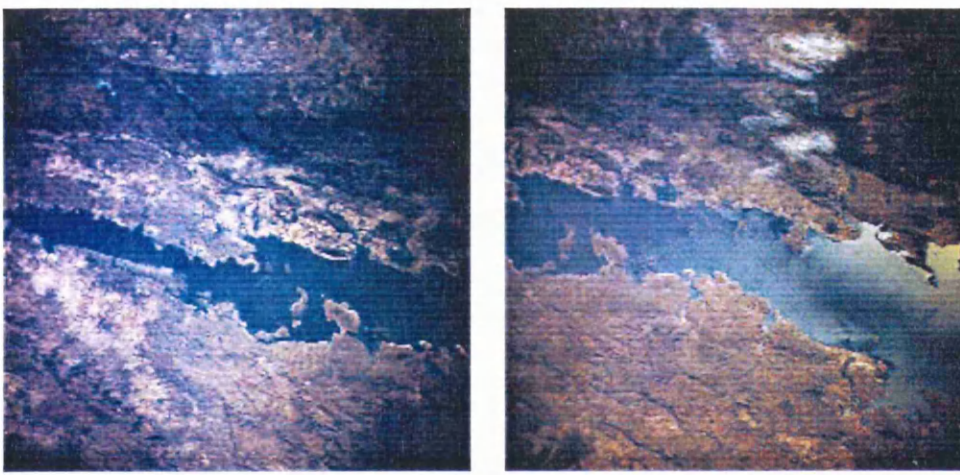


Figure 2.4 Space Shuttle handheld camera images of Lake Kariba (© NASA)

The two images, acquired a few moments apart, have about 40% overlap. Since no supplementary information (such as camera orientations) is available it is not possible to use photogrammetric techniques to mosaic these images. Therefore image registration has to be used. A total of seven tie points were selected in the area of overlap between the two images. These tie points were used to determine the parameters of an affine transformation, which in turn was used to register the image on the right to the image on the left. The resulting mosaic is shown in figure 2.5.

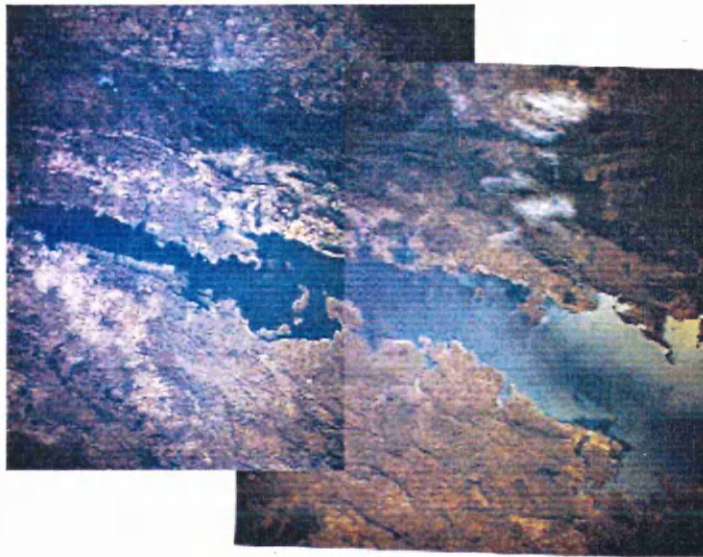


Figure 2.5 Mosaic created using affine transformation

Figure 2.5 shows that the registration has been fairly successful, and this is borne out by the RMS residual of the tie points: the RMS residual in the x direction is 0.55 pixels, whilst the RMS residual in the y direction is 0.42 pixels. If the registration is repeated using a second order (quadratic) transformation function, the tie point residuals can be dramatically improved: the RMS residual in the x direction becomes 0.22 pixels, whilst the RMS residual in the y direction becomes 0.01 pixels. The mosaic created using the quadratic transformation function is shown in figure 2.6.

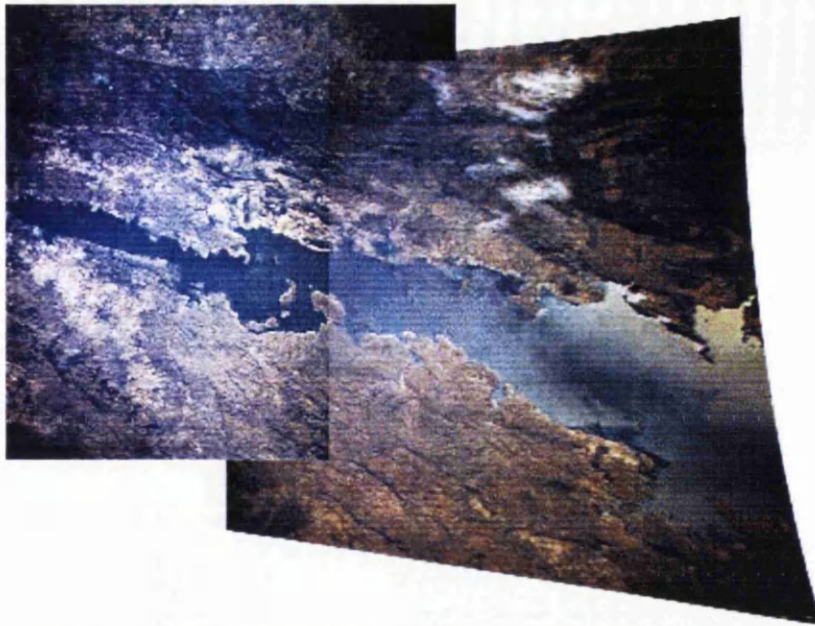


Figure 2.6 Mosaic created using quadratic transformation

It can be seen from figure 2.6 that even though the tie point residuals are far better, the resulting mosaic looks very different to the previous one. Check point residuals spread across the images would reveal that the registration was in fact far worse – contrary to what is suggested by the tie point residuals.

This example also illustrates well the importance of selecting the correct transformation function. Using exactly the same tie points, but a different transformation functions leads to a very different result. An obvious point which is very clear is the rate of degradation of the higher order polynomials. In figure 2.6, a close examination of the area of overlap shows the registration to be very good (better than that in figure 2.5), but as the distance from the tie points increases, the quality of the registration degrades very fast. Higher order polynomial transformations (third and fourth order) would show this effect even more. For lower order polynomials (first and zero order) this is not the case (see figure 2.5). This can also be illustrated with a one dimensional curve fitting example. Figure 2.7 shows three curves. The red curve represents the function:

$$y = \frac{1}{1 + 25x^2} \quad (2.9)$$

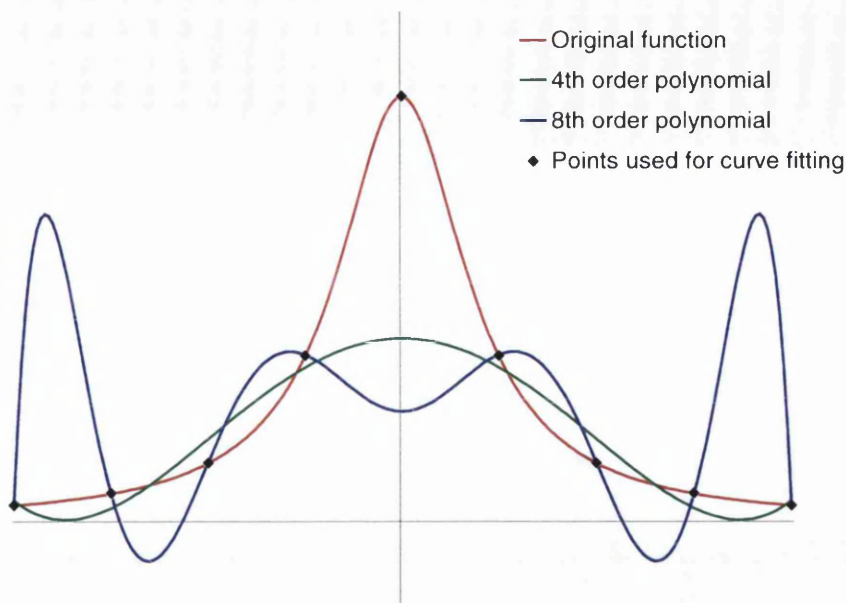


Figure 2.7 Non-linear function (red curve) with 4th and 8th order best fit polynomials (green and blue curves respectively)

This curve has been approximated by two ‘best fit’ polynomial curves: a 4th order polynomial (shown in green) and an 8th order polynomial (shown in blue). The parameters of these two polynomials have been generated by using a least squares fit to the nine black points on the curve. The 8th order polynomial gives the best fit to the data points, shown by the fact that the blue curve passes exactly through eight of the nine points. However, it can be seen that this curve does not represent the original curve very well at all, since between the points it diverges from the original curve quite considerably. Conversely, the 4th order polynomial represents the original curve much better, but it does not pass through the points nearly as accurately. If only the residuals at the points were considered as a measure of accuracy of fit, the 8th order curve would appear to fit much better than the 4th order curve, but by looking at the regions between the points, it can be seen that the 4th curve fits much better. This reiterates the need for independent check points when assessing the accuracy of fit.

It can also be seen from figure 2.7 that neither of the two best fit polynomials fit the original very well at all. With image registration it is generally accepted that high order polynomials will not improve registration accuracy - in fact anything above a second order polynomial will probably give a less accurate result than anything below a second order. High order polynomials should only be used in image registration when it is clear that a polynomial function is causing the distortion, and its order is known. Thus there exists an interesting dichotomy: low order polynomials do not necessarily register images well, but increasing the order of the polynomials only makes things worse. A

possible solution is to use functions other than polynomials to correct the distortions, or to use the polynomials in a more ingenious way.

The similarity and affine transformation functions described above are very useful for registering images which are acquired from a similar viewpoint and have a scene-to-sensor distance which is much greater than variations of terrain in the scene. It is when these two points are not satisfied that higher order polynomials are used. However, if the terrain is flat, but the sensor locations are very different, the perspective transformation will model the geometry well (Peet, 1975). The perspective transformation function can be expressed as:

$$X = \frac{(ax + by + c)}{(dx + ey + 1)} \quad (2.10a)$$

$$Y = \frac{(fx + gy + h)}{(dx + ey + 1)} \quad (2.10b)$$

Although parallelism and angles are not preserved in the transformation, straight lines remain straight. This function requires a minimum of four tie points to evaluate the parameters of the transformation.

All of the transformation functions described above are *global* transformation functions, meaning they perform the same transformation on the entire image. The other group of transformation functions are *local* transformation functions, which apply a slightly different transformation function to different parts of the image. Local transformation functions can be further split into two groups: surface fitting functions and piecewise polynomials. Both of these are described below.

Surface fitting functions are functions which can be used to fit a surface to a set of points. Consider the set of tie points $\{(x_i, y_i), (X_i, Y_i) : i=1, \dots, n\}$. These can be rearranged into two sets:

$$\{(x_i, y_i, X_i) : i=1, \dots, n\} \quad (2.11a)$$

$$\{(x_i, y_i, Y_i) : i=1, \dots, n\} \quad (2.11b)$$

Two surface functions f_x and f_y can now be defined which interpolate the sets of 3D points given above. In the special case of the affine transformation (equations 2.6a and 2.6b), these surfaces are planes. Three methods of interpolating the surfaces from the

sets of points are surface splines, multiquadratics, and sum of Gaussians. Surface splines are defined by the function:

$$f(x, y) = a_0 + a_1x + a_2y + \sum_{i=1}^n F_i d_i^2 \ln d_i^2 \quad (2.12)$$

The parameters of this function are determined by using a set of tie points to generate a set of simultaneous equations, as with the polynomial transformation function, and by setting specific boundary conditions. Multiquadratics are defined by the function:

$$f(x, y) = \sum_{i=1}^n \alpha_i \sqrt{(x - x_i)^2 + (y - y_i)^2 + r^2} \quad (2.13)$$

Once again, a set of tie points are used to generate a set of simultaneous equations in order to solve for the parameters α_i . The parameter r , which controls the smoothness of the surface, has to be set manually. Sum of Gaussian surfaces are defined by the function:

$$f(x, y) = \sum_{i=1}^n V_i G_i(x, y) \quad (2.14)$$

As with the two previous examples, a set of tie points are used to generate a set of simultaneous equations, which in turn are used to solve for the parameters V_i . As with the parameter r in the multiquadratic function, σ gives a measure of the smoothness of the surface, and has to be set manually.

Although the above three functions are very good at interpolating a smooth surface through a set of points, there are some problems with applying this to image registration. Firstly, the number of tie points has to be relatively small (less than fifty). For each pair of tie points, two simultaneous equations are generated. All these equations have to be solved to determine the parameters of the transformation function, but this becomes computationally impractical for large numbers of equations. Secondly, since the surface is fitted to all the tie points it is very difficult to determine if there are any blunders in the set of tie points by looking at the transformation function alone. Both of these issues have been addressed by Satter and Goshtasby (1997), who suggest using an approximation rather than an interpolation for the surface fitting (hence reducing the number of equations to be solved), and displaying the transformation functions to highlight any blunders. A final problem of using surface fitting functions as

transformation functions in image registration is that they do not really reflect the physical situation. Therefore care must be taken when using these types of functions, and checks should be built into the processing to ensure the result is accurate. This is similar to the situation with high order polynomials - the functions may provide an apparently better fit, but care must be taken to ensure this is so.

The second group of local transformation functions are piecewise polynomials. As the name suggests, different transformation polynomials are applied to different parts of the image in the registration process. For example, if two images that are being registered feature an area of flat terrain and an area of mountains, then a low order polynomial would be used to register the flat areas, and a higher order polynomial or a surface fitting function would be used to register the mountainous areas. The image is effectively split up into a number of sub-regions, each of which have their own transformation function. Although this may seem like a reasonably good way of registering two images with varying distortions which cannot be modelled explicitly, there are a number of issues which have to be addressed, such as:

- ❑ the strategy for splitting up the image into separate tiles to be registered;
- ❑ the selection of the transformation function for each of those tiles; and
- ❑ the method dealing with boundaries between adjacent tiles.

The solutions to these issues are very much dependent upon the type of imagery being registered, and therefore cannot be solved in general. Other problems also exist, such as the large number of tie points required, the large computation time and the increased validation of the transformation functions.

2.3.4 Calculation of transformation parameters

In a study of rectification accuracy, Mather (1995) describes two algorithms (matrix inversion and orthogonal polynomials) for determining the parameters of second order polynomial transformation functions. It is not clear which of these algorithms was used for determining the parameters in the examples presented in this thesis since that information was not supplied with the software. However, this was not considered to be an important issue since the same software was used throughout this thesis, so the results presented in the later chapters are consistent with each other.

2.3.5 Resampling

The last step in the image registration procedure is the resampling of the slave image to the master image. Image resampling involves the reformation of an image to a new grid, and comes about because a digital image has pixel values at discrete locations - the image does not represent a continuous distribution of values. Figure 2.8 shows two grids overlaid on top of each other. The master image is represented by the black grid, and the slave image is represented by the blue grid. Pixel values in the slave image (the blue grid) need to be expressed in the co-ordinate system of the master image, the black grid. A number of different resampling schemes can be used to determine the pixel value of the grey shaded square in the master grid. These include:

- ❑ nearest neighbour;
- ❑ bilinear interpolation; and
- ❑ cubic convolution.

With nearest neighbour resampling, the nearest pixel in the slave image is assigned to the grey square in the master grid. In figure 2.8, the nearest pixel is denoted by the letter *a*. Bilinear interpolation interpolates a value for the pixel using the nearest four pixel values in the slave image, represented by the letters *a* and *b*. Cubic convolution uses the nearest sixteen pixels to determine the new value for the pixel in the master grid (represented by the letters *a*, *b*, *c*).

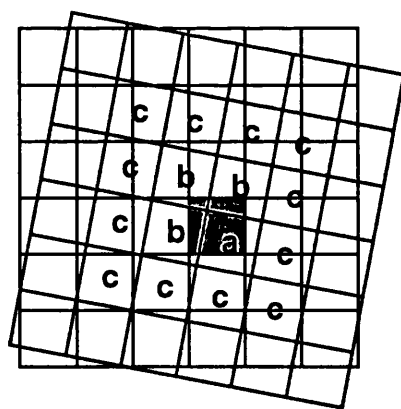


Figure 2.8 Resampling schemes

Nearest neighbour resampling is the quickest to perform of the three schemes, but generally gives the coarsest result. Cubic convolution usually gives the smoothest result, but is most computationally expensive. In the past, bilinear interpolation was seen as a good balance between quality of result and computing time. However, the rapidly increasing power of modern computers means that CPU time is not really a major processing problem, so cubic convolution can be seen as the best resampling method.

2.3.6 Conclusions

This section has described the four main steps of manual image registration. The next section illustrates how some of these steps can be automated, in order to remove the human operator from the processing chain. It also illustrates some of the problems associated with automatic image registration, in addition to presenting a review of previous research.

2.4 Automatic image registration

2.4.1 Introduction

It is generally accepted that there are two methods which can be used to match portions of an image automatically (Heipke, 1996):

- ❑ area based matching; and
- ❑ feature based matching.

Area based matching works by extracting a small area (image chip) from the master image and scanning the slave image in order to find a similar area (Gruen, 1986; Otto and Chau, 1989). Areas are matched by maximizing a cross correlation function. Although very successful in certain situations, this method suffers from a number of limitations. The images to be matched have to be radiometrically very similar, preferably imaged by the same sensor. In addition, the camera model for the sensor needs to be known in order to ensure the images have minimal differences in scale and azimuth. Furthermore, the success of matching can depend on external influences, such as lighting conditions and cloud contamination. For these reasons area based matching is, in general, only applicable to single source image registration. Day and Muller (1989) describe a procedure which successfully registered a pair of SPOT images using area based matching. This procedure can also be used to generate DEMs from a stereo pair of images, and therefore orthorectify those images, when ground control data is available. However, since it is difficult to apply area based matching to multisensor data sets, the remainder of this section describes techniques for feature based matching.

As with manual image registration, there are four main steps that make up automatic feature based image registration (Brown, 1992). These are:

1. extraction of primitives from each image;
2. matching of corresponding primitives from each image;
3. determination of the transformation function; and

4. transformation of the image.

Steps 1 and 2 of the feature based registration procedure (extraction and matching of primitives) are analogous to step 1 of the manual registration procedure (selection of tie points). With manual image registration the tie points are identified in one step, but automation of this task requires it to be split into two distinct processes. The first process is the extraction of a large number of features (known as primitives), and the second is the matching of those primitives to give a useful set of tie points. Step 3 of the feature based registration procedure (determination of the transformation function) is analogous to steps 2 and 3 of the manual registration procedure (selection of the transformation function and determination of the parameters of that function). In the automatic procedure, the decision of which transformation function to use has to be made automatically. This is one of the more difficult tasks to automate because there is not necessarily a right or wrong answer. The final steps of both the automatic and manual registration procedures are the same as each other: registration of the slave image to the master image, using a the transformation function and a resampling scheme. The next three sections describe these steps in more detail.

2.4.2 Extraction of primitives

The types of primitives to be used in the registration procedure are selected by examining the images to be registered, and considering any supplementary data that may be available (such as topographic data sets) which may help the extraction. It is not necessarily possible to extract all types of primitives from a particular image, so in this processing step a decision has to be made as to which primitives are used, based on the image type and the image content. It is important that the same primitive is used for each image that is being registered since it is usually not possible to match one type of primitive extracted from one image with a different type extracted from the second image (for example, it is not possible to match patches with point features). However, there is no need to limit the matching to only one type of primitive. If both linear features and regions can be extracted from the images which are being registered, then both sets of primitives can be used.

Supplementary data is very useful for improving the accuracy and efficiency of feature extraction algorithms since it can be useful for highlighting particular types of primitives. For example, a topographic map of field boundaries can help to segment an optical image of an agricultural area (Li and Muller, 1991).

2.4.3 Matching of primitives

The matching process takes the primitives that have been extracted from the images and generates from these a set of tie points (i.e. points which can be identified in both images). There has to be a one to one correspondence between tie points, so each point in the master image can only be matched with one point in the slave image, and *vice versa*. Primitives are matched based on some or all of their attributes, such as size, shape or radiometry. When two corresponding primitives are matched they may generate one pair of tie points, or numerous pairs, depending on the type of matching algorithm used. For example, if a pair of regions are matched, a single pair of tie points could be generated from the centroids of the regions, or numerous pairs could be generated from the pixels which make up the edges of the regions. Figure 2.9 shows two matched polygons (in grey). A single tie point could be defined, based on the centroids of the polygons (blue diamonds) or multiple tie points could be defined, based on salient points on the perimeters on the polygons (red crosses).

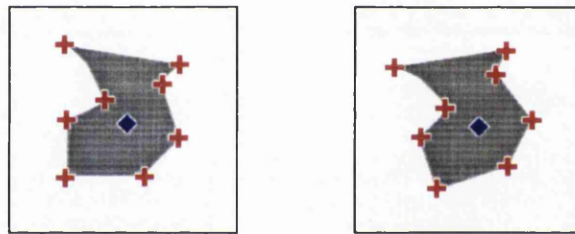


Figure 2.9 Tie point extraction from matched polygons

2.4.4 Determination of the transformation function

Selection of the transformation function depends on all the data available: image types, sensor geometry, match point distribution, supplementary data etc. When the transformation function has been selected its parameters are determined using the tie points extracted from the images. The final steps of the automatic registration procedure (image transformation and resampling) can easily be performed automatically and requires no further discussion here.

2.4.5 Review of automatic registration techniques

A recent survey of current image registration techniques was given by Fonseca and Manjunath (1996), which covered area and feature based matching in both the spatial and frequency domain. An earlier survey by Brown (1992) was far more wide-ranging and compared numerous different applications of image registration, including machine vision and medical imaging. A summary of registration techniques applicable to the photogrammetric processes of interior, relative and absolute orientation were given by Heipke (1996). This section will be limited to reviewing relevant research on feature

based matching only. The research described in this section is listed under author in alphabetical order.

Abbasi-Dezfouli and Freeman (1994)

Abbasi-Dezfouli and Freeman (1994) used patches as matching primitives to register SPOT stereo pairs. The aim was to register the images fully automatically and determine the terrain height across the region of interest. The primitives were extracted by searching the image for patches of uniform colour and matched using several criteria corresponding to the patch, such as area, dimensions of bounding rectangle, perimeter, linearity, concavity and relative geometry. After the patches had been matched, tie points were generated by matching significant points on the boundaries of corresponding patches with each other using a correlation method. Results showed that tie points were generated from both correctly and incorrectly matched patches, leading to a tie point file containing some false points. No mention is made of how the match points would be used to register the images and create a terrain model. The test images used were of the order of 500x500 pixels.

Dowman, Boardman and Newton (1996)

The method described by Dowman et al. (1996) aims to register large same sensor images (SPOT stereo pairs) fully automatically. The matching primitives are significant points that are extracted from the images using the Förstner interest operator (Förstner and Gulch, 1987). An approximate initial affine transformation is used to register the images. The interest points are projected from the slave image into the master image to be matched. Generally a one to one match can be made based on location in the image alone, but where the match is ambiguous a cross correlation is applied to an area around the points, the result of which is used to select or reject a match. After all possible matches have been completed a final test is carried out with intensity based matching. The resulting tie points are used to generate the parameters of an affine transformation, which in turn is used to register the images. For large images a pyramid structure is used. Results showed an RMS error of 2 pixels (approximately 20m on the ground) from images of 1024x1024 pixels.

Flusser and Suk (1994)

Flusser and Suk (1994) used patches as matching primitives to register SPOT and Landsat TM images. The primitives were extracted using a segmentation technique that worked by applying an edge detection algorithm to the image and thresholding the

result. Patches with perimeter lengths between 10 and 100 pixels were extracted before being refined using a contour tracing technique in the original image. The patches were represented by affine moment invariants (AMIs), which are moment based descriptors of planar shapes invariant under the affine transformation. The patches were matched by comparing the AMIs of all the patches in both images, and selecting the three best matches. The co-ordinates of the centroids of these matched patches gave three tie points that were used to determine the parameters of an affine transformation. The slave image was then registered to the master image and a region to region correspondence was established using a nearest neighbour scheme. When all patches had been matched, the co-ordinates of all the conjugate patches were used to generate tie points, which in turn were used to generate the parameters of an affine transformation, based on a least squares method.

Flusser (1992)

The technique described by Flusser (1992) attempts to simplify the processing problems associated with surface splines (see Goshtasby (1988) below) by using an adaptive mapping (AM) algorithm. The number of calculations needed to solve the two surface spline functions increases very quickly with increasing numbers of tie points and larger images. The AM algorithm splits the image up into smaller tiles and uses a much simpler equation to represent the surface spline function. Four different types of equation were tested on two different images and sub-pixel accuracy was achieved each time. The method relies on a number of correctly identified tie points being selected in advance.

Goshtasby (1988)

Goshtasby (1988) describes a technique for registering images with geometric distortions using two surface splines which represent the x and y components of a mapping function (see § 2.3.3 above). The method relies on a number of correctly identified tie points to solve the two equations. Examples of registering a Landsat MSS scene to a simulated TM image gave an RMS error of 0.63 pixels, compared to 1.9 pixels when a polynomial and least squares techniques was used. This technique does have its drawbacks: it cannot be used for a large number of tie points (in this case large means greater than 50), and the mathematical model does not reflect the physical situation very well.

Goshtasby, Stockman and Page (1986)

Goshtasby et al. (1986) proposed a registration technique where patches were extracted as matching primitives using image segmentation. Corresponding patches were matched with each other using a clustering technique involving the centroids of the patches (see Stockman et al., 1982 below). The result of the clustering is an approximation of the scale factor, rotation and translation required to register the images. By registering the images, corresponding patches can be identified and their edges refined so that they become optimally similar. The centroids of the patches are then re-determined to an anticipated higher degree of accuracy. The co-ordinates of these centroids are then used to determine the parameters of a polynomial transformation, which is in turn used to register the images. Sub-pixel accuracy was achieved when this procedure was tested on Landsat MSS and simulated TM data.

Li, Manjunath and Mitra (1995)

Li et al. (1995) proposed two similar methods for registering (i) Landsat and SPOT images, and (ii) Landsat and Seasat SAR images. The first method is a contour matching scheme where both open and closed contours are extracted from the pair of images as matching primitives. Closed contours are matched by comparing attributes such as chain codes and invariant moments. Open contours are matched by detecting salient points such as corners and using the contour segments around these points as templates that are matched between the images. Centroids of closed contours and salient points of open contours were used as tie points. The parameters of an affine transformation were determined from the tie points using a least squares technique.

The second method for registering optical to SAR data is based on an active contour model, where the edges extracted from optical data are used to help locate the edges in the SAR data. To do this the images have to be pre-registered to an accuracy of approximately 5 pixels, which can be done by either using ephemeris data or manually selected control points. When the contours have been extracted from the SAR image, the registration algorithm proceeds as described above. An RMS error of 1 to 2 pixels was achieved for a selection of test images.

Morgado and Dowman (1997)

Morgado and Dowman (1997) registered aerial photographs to a map using patches as the matching primitives. A coarse registration is generated by matching patches on the basis of attributes such as size, shape and perimeter length. The coarse registration is

then refined using a technique based on dynamic programming. Tie points are generated from the pixels that make up the edges of the patches. The image is registered to the map using an affine transformation, the parameters of which were determined from the tie points using a least squares technique. An RMS error of approximately 7m (2 pixels) in the x - y plane and 1.2m in the z direction were obtained. This method is very similar to that used by Newton et al. (1994).

Stockman, Kopstein and Benett (1982)

The method described by Stockman et al. (1982) uses a clustering technique to generate tie points. Point features are extracted using straight edge detectors, curved edge detectors, circle detectors and intersection detectors, and paired up to form vectors. These vectors are used as the matching primitives. Each primitive from the master image is matched with a primitive from the slave image and the rotation, scale and translation components of the transformation required to register these primitives are determined. This is repeated for all possible combinations of master and slave primitives and the results are plotted in rotation-scale-translation (RST) space. By identifying clusters of points in RST space it is possible to determine the parameters of the transformation required to register the slave image to the master image.

Tseng, Tzen, Tang and Lin (1997)

The method described by Tseng et al. (1997) matches the closed boundaries of regions to determine the parameters of the transformation function. Areas of homogeneous grey level which are greater than a specific size are extracted from the images, after which their boundaries are encoded using Fourier descriptors. The boundaries can then be matched in the frequency domain and the relative orientation of the images can be determined. The system was implemented using a neural network.

Ventura, Rampini, Schettini (1990)

Ventura et al. (1990) describe a method of automatic image registration using attributes of structures (ellipticity, inclination and thinness) extracted from the images which are invariant under scaling, rotation and translation. The structures are extracted from the image by applying a segmentation algorithm after which an algorithm based on fuzzy logic is used to compare and match similar patches. Tests with Landsat TM and SPOT data gave an RMS error of 2.5 pixels.

This completes the review of automatic image registration techniques. The next section introduces some problems associated with automatic image registration, but before this, table 2.1 summarizes the work described above.

Authors	Matching primitive	Matching technique	Transformation function
Abbasi-Dezfouli & Freeman (1994)	Patches of uniform grey level	Comparison of size, shape, and relative geometry	Affine
Dowman et al. (1996)	Interest points	Relative geometry and cross correlation	Affine
Flusser and Suk (1994)	Patches	Comparison of AMI and relative geometry	Affine
Flusser (1992)	Manually selected points	Manual matching	Simplified surface splines
Goshtasby (1988)	Patches	Clustering in RST space	Surface splines
Goshtasby et al. (1986)	Patches	Clustering in RST space	Polynomial function
Li et al. (1995)	Open and closed contours	Comparison of invariant moments and chain codes	Affine
Morgado and Dowman (1997)	Patches	Comparison of size, shape, and relative geometry	Affine
Stockman et al. (1982)	Vectors connecting interest points	Clustering in RST space	Similarity
Tseng et al. (1997)	Fourier transform of closed boundary	Comparison of Fourier transforms	Affine
Ventura et al. (1990)	Patches	Comparison of ellipticity, inclination, thinness	Affine

Table 2.1 Summary of automatic registration techniques

2.4.6 Difficulties associated with automatic image registration

The previous sections have given a description of how images can be automatically registered to each other and presented some of the current techniques available, but it should be remembered that there are still a number of problems which need to be overcome before a fully automatic and accurate image registration system becomes realistic. Three problems which are discussed in this section are:

- ❑ the difficulty of recognizing features acquired from multiple sensors;
- ❑ application of image registration algorithms to numerous different types of images; and,
- ❑ the need for an intelligent system to select which algorithms to use in particular situations.

The problem of recognizing similarities between pairs of images is fundamental to automatic image registration and perhaps the most difficult problem to solve. The problem is twofold: firstly, if the approximate alignment of the images is not known then the spatial distribution of features will be different, and secondly, if the images are acquired by different sensors then corresponding features will have different radiometric properties. These two aspects of the problem of feature recognition are illustrated with some examples below.

Figure 2.10 shows two images that correspond to the same region in Death Valley, USA, but which are not aligned to each other. The image on the left was acquired by the SIR-C/X-SAR sensor whilst the image on the right was acquired by the SPOT sensor. Although the images show the same region it is impossible to recognise any common features with any degree of certainty. If it is not possible for the human visual system to recognise features in pairs of images, then it will be extremely difficult to develop a computer algorithm which can do this automatically. If the images were approximately aligned, then the possibilities of recognizing common features is greatly increased, but if the images cannot be approximately aligned, then it may not be possible to use automatic image registration algorithm based on feature matching.

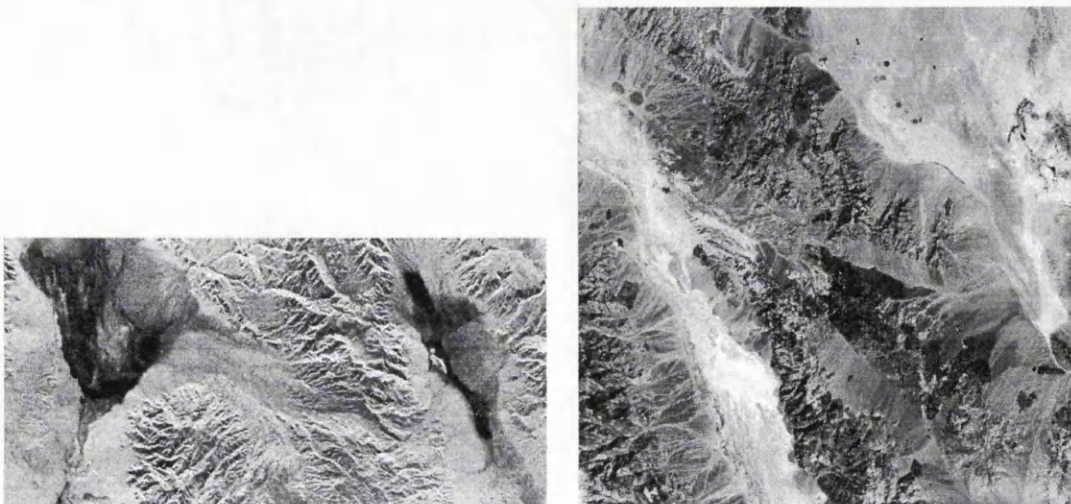


Figure 2.10 Radar and optical images of Death Valley, USA (© NASA)

Figure 2.11 shows two images of the same region in Southern France. The image on the left is a SPOT panchromatic image and the image on the right is an ERS-1 SAR image. In this example there are one or two features that can be recognised in both images, but since the images have different projections the corresponding features have different sizes and shapes. Furthermore, since the images were acquired using different sensors the grey values of common features are also quite different. Therefore, although in this case it is possible for the human visual system to recognise common features it is still very challenging to develop computer algorithms that can perform a similar task.

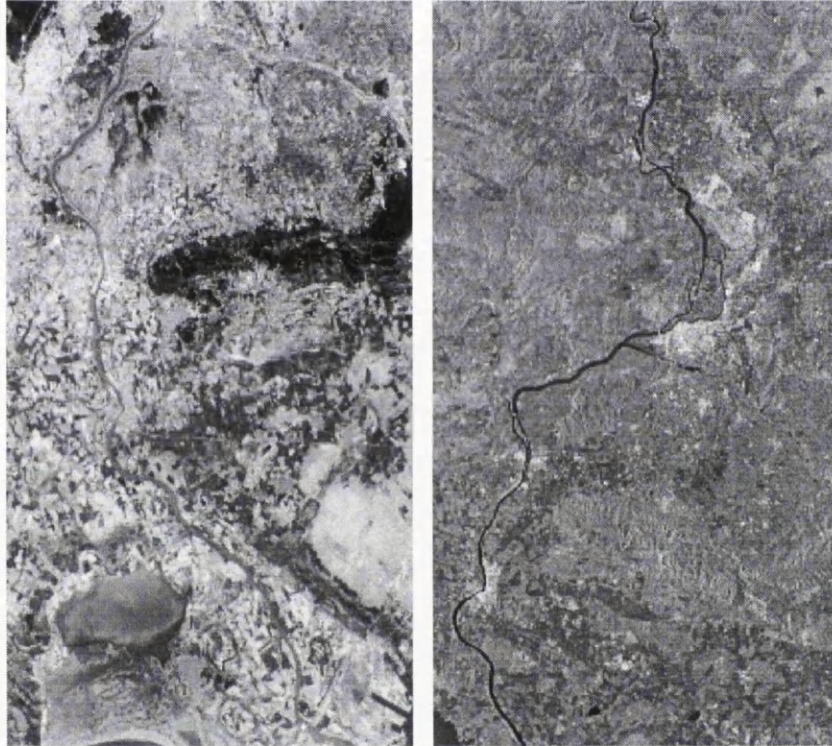


Figure 2.11 Radar and optical images of Southern France

Even when a pair of images are aligned fairly well it can still be difficult to recognise common features. Figure 2.12 shows regions extracted from the above pair of images of Southern France. The figure on the left is a SPOT panchromatic image with a pixel size of approximately 10m whilst the figure on the right is an ERS-1 SAR image with a pixel size of approximately 12m. The SPOT image has been registered to the SAR image using four manually selected tie points. Even though the images have the same projection and cover the same area, there are few features that can clearly be identified in both images. Once again, if the human visual system cannot recognise pairs of corresponding features, then developing a computer algorithm to perform this task will be very difficult.

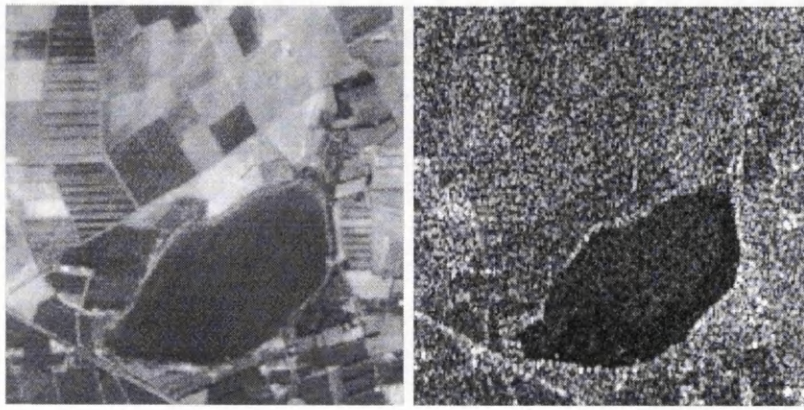


Figure 2.12 Registered radar and optical images of Southern France (*R and L resp.*)

The pairs of images shown above highlight some of the difficulties of matching features in pairs of multisensor images. The problems persist at both small scales (for approximate alignment of large images) and large scales (for accurate registration of smaller scenes).

The second difficulty introduced above is that of global application of algorithms. Many of the automatic image registration algorithms described in § 2.4.5 perform very well with the test data sets provided by the authors, but often these algorithms will fail when used with different types of data sets. The problem that has to be overcome is the development of algorithms that can be applied to almost any data set and still produce useful and accurate results. To do this, the automatic image registration system will have to be able to solve practical problems, such as multisensor feature recognition described above, as well as more fundamental problems, such as intelligent decision making, described below.

The third difficulty mentioned above is that the method of automatically registering two images is not well defined. The review of recent image registration research in § 2.4.5 above shows that there are a number of different algorithms which can be used at each stage in the processing, and it is often far from clear which is the best one to use. A challenge to researchers of automatic image registration is the development of an intelligent system that can make wise choices as to which algorithms to use and when. Until the implementation of such a system is realised, there will always be the need for human intervention in automatic image registration software to ensure a satisfactory result.

2.4.7 Conclusions

This section has described in detail the subject of automatic image registration – it has been presented in some depth since it is fundamental to the work of this thesis. The

initial sections introduced the theory of automatic image registration and described how it can be implemented using a four stage procedure. This procedure has been well documented by a number of researchers (Brown, 1992; Fonseca and Manjunath, 1996) and is generally accepted as the standard approach to take for automatic image registration. However, the actual algorithms which can be used at each stage of the procedure can vary considerably; a number of different combinations of algorithms were presented in the review in § 2.4.5. Some of the authors of the work described in this section claimed to achieve certain accuracies in terms of RMS errors, but without knowing exactly how they arrived at their results, it is difficult to know how reliable these claims actually are. Therefore, it seems more reasonable to judge the work on the actual methods presented rather than on the results. The range of matching primitives employed is extremely varied, as are the techniques for matching these primitives. Most of the methods described use polynomial transformation functions to register the pair of images, except for Flusser (1992) and Goshtasby (1988), whose transformation functions based on surface splines. It is not possible to select particular methods as being any better than any of the others, but the comparison of techniques does illustrate well how varied image registration algorithms can be.

The final section highlighted three difficulties associated with automatic image registration: recognition of features in dissimilar images, application of algorithms to a wide range of images, and choice of algorithms. It may not be possible to resolve these difficulties with current techniques, but these are issues that must be taken into consideration in the development of an automatic image registration system.

Chapter 1 introduced image rectification: the registration of an image to a map, resulting in an image which can be expressed in ground co-ordinates, but with no correction made for terrain and sensor induced distortions. Automatic image rectification has not been described in detail in this thesis, but it deserves a mention since it is very similar to automatic image registration. A lot of the techniques associated with automatic rectification are the same as those for automatic registration: images are rectified by extracting features from the image and the map and matching them with each other. However, automatic rectification experiences other problems which are not encountered in automatic registration, such as those caused by generalisation of features in the map data, or the recognition of corresponding features in both the image and the map. More details about automatic rectification of aerial photography and satellite data can be

found in Dowman (1998), Morgado and Dowman (1997), Vohra and Dowman (1997) and Vohra et al. (1996).

This concludes the discussion of automatic image registration. The next section explains how the work presented in this thesis attempts to solve some of the problems associated with the automatic registration of SAR and SPOT images.

2.5 Proposal of this thesis

2.5.1 Introduction

The first chapter of this thesis introduced the concept of data integration and explained its importance. This chapter has described in more detail the issues associated with the geometric aspects of data integration. This section will now describe how the work presented in the remainder of this thesis proposes to solve some of the problems encountered with the automatic registration of multisource remotely sensed images. The overall aim of this study is the development of a fully automatic system for registering SAR and SPOT images. However, this would be too large an undertaking for a study of this nature, so therefore only a few key elements of such a system have been researched. Before the details of these key elements are discussed, the concept of an "ideal" automatic image registration scheme is proposed.

2.5.2 The "ideal" automatic image registration system

The basic theory of automatic image registration has now been described. Using this structure it is possible to define an ideal automatic image registration system. Many other researchers have also proposed ideas for automatic image registration systems, such as Dowman (1998), Dowman et al. (1996), Le Moigne et al. (1997) and Rignot et al. (1991), but the concept presented here takes into account factors such as robustness, reliability and practicality of implementation. Very often these factors are ignored in favour of other details such as accuracy of registration results.

Although not described in the processing chain presented in § 2.4.1, the first step in any automatic image registration system is the preprocessing step, which is concerned with preparing the images for registration. The procedures involved include:

- importing and formatting the data:
 - reading the images; reading the ephemeris data; reading the supplementary data; converting all data to a standard format.
- preprocessing the data:

approximate alignment; geocoding and resampling (if necessary).

- user validation.

The ideal system would have to be able to recognise and read numerous data formats. The preprocessing step would convert all of the data to a standard format and then ensure it is suitable for feature based matching. That suitability will depend on the percentage overlap and the content of the images. Approximate alignment calculated from the ephemeris data can be used to determine the percentage overlap, but it is much more difficult to automatically assess the suitability of the image content for matching. No suggestions for performing this task are offered here.

Next, the supplementary data sets must be considered, which can include DEMs, topographic data, GCPs or ephemeris data. Using this supplementary data, and knowledge of coverage of the images, it must be decided if any preprocessing is possible, such as geocoding, resampling to similar grid size, or performing any geometric corrections. If very little is known about the images, or there is no supplementary data, then the preprocessing will be rather limited. It would be desirable, however, to carry out as much preprocessing as possible in order to increase the success of the feature extraction and feature matching algorithms.

The final stage of the preprocessing is the user validation. The system would provide the user with details of all the tasks performed to ensure it was prudent to continue to the next step of the registration.

In the primitive extraction step of the processing, as many primitives need to be extracted as possible, but an automatic validation step would have to be included to filter out as many false primitives as possible. At this stage, supplementary data such as topographic data or change detection information would be very useful. Topographic data could be used as *a priori* knowledge for high level feature extraction algorithms, and change detection data would help guide those algorithms away from areas in the image where changes may confuse the matching process. An automatic test of the quantity and distribution of the primitives across the image would show whether the extraction process had been successful or not. If it was found that too few primitives had been extracted from the images, then the process of primitive extraction would have to be repeated.

As with the previous step, the matching step would be greatly improved by incorporating a validation procedure. This is required to ensure there are a large number of match points with a good distribution across the whole image. If it appears there are insufficient match points, it may be necessary to return to the previous stage and extract more matching primitives. It may be possible to improve the matching of primitives by using the matching results generated at this stage to align the images more closely. Hence, an iterative procedure could be set up between primitive extraction and matching. It may also be possible to generate information on changes that may have taken place between the acquisition of the images from the match point data. This could also be used as a guide for extracting primitives in the previous stage if it were used in an iterative structure.

The ideal automatic image registration system would have to be able to select the most appropriate transformation function from a library of functions. Although selection would be aided by analysis of the supplementary data, it is likely that a number of different functions would have to be tested to see which would give the best result.

Figure 2.13 shows a flow diagram of the ideal automatic image registration system. The important points to highlight are that there are numerous validation steps and feedback loops, ensuring that errors should not be propagated through the system, and there is plenty of opportunity for the system to adapt itself to the data being used. If no match points can be found using one particular set of algorithms, then the system recognizes this, and repeats the process with a different set of algorithms, until a final result is obtained. The key is to utilize libraries containing many different algorithms, therefore guaranteeing a large redundancy of processing techniques.

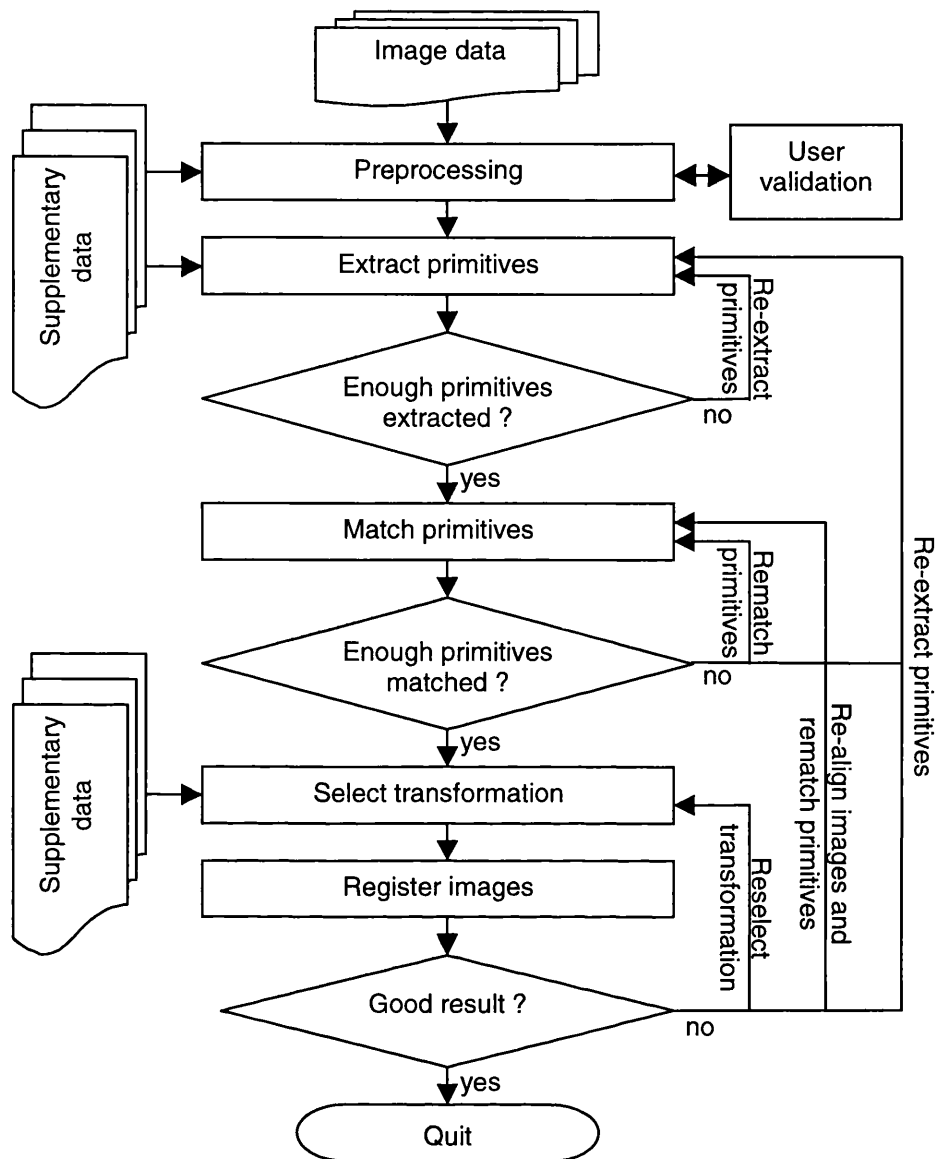


Figure 2.13 The ideal automatic image registration system

This section has outlined the components of an ideal automatic image registration system. The next section explains which of these components are researched in detail in this thesis.

2.5.3 Elements of automatic registration researched in detail

The previous section has illustrated that there are a number of aspects of automatic image registration that need to be researched. Firstly, the specific algorithms for performing tasks at each stage of the processing need to be developed, and secondly, those algorithms need to be combined together into a single automatic system. It seems reasonable that the development of specific algorithms should take precedence over the development of the system as a whole. It was therefore decided that this thesis would investigate and develop algorithms for automatically registering SAR and SPOT images based on patch matching, with emphasis placed on the extraction and matching steps in

the processing chain. After an examination of different feature based image registration techniques, it was decided that the processing chain depicted in figure 2.14 would be researched.

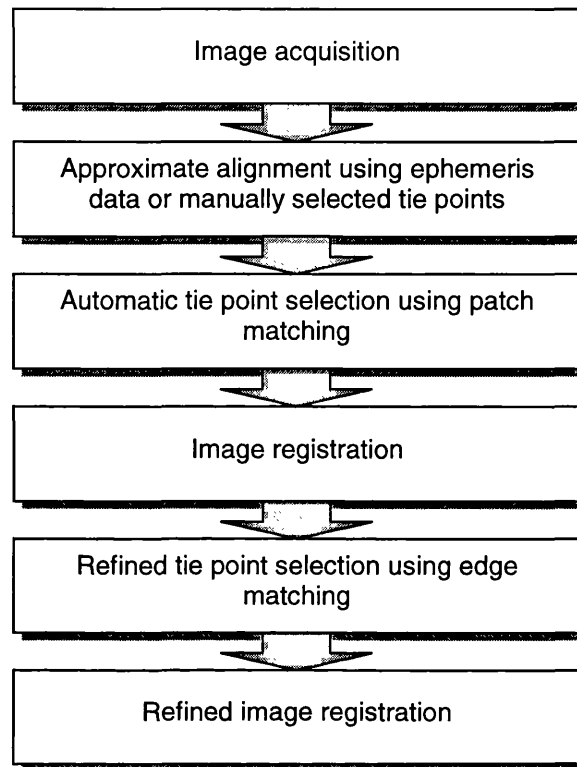


Figure 2.14 Flow chart of proposed image registration procedure based on patch and edge matching

The proposed registration procedure consists of three registration steps. The first step is an approximate alignment of the images using either a few manually selected tie points, or tie points determined from the ephemeris data. The precision of the registration does not have to be high at this stage – misalignments of up to 100 pixels are acceptable at this point in the processing. However, it is important that large rotation and scale differences are eliminated so that the second stage of the processing chain, the automatic selection of tie points using patch matching, is able to locate corresponding features in the pair of images. As a result of matching corresponding patches it should be possible to register the images to a precision of less than a few pixels (depending on the number and distribution of matched patches and the choice of polynomial transformation function). The final stage in the processing uses edge matching to refine the registration results. An edge matching algorithm is employed which can locate a large number of common points between a pair of images, but in order to do so, the images have to be registered to a precision of less than a few pixels. This registration procedure relies on each processing step being able to align the images more precisely than the previous one, and then passing its results on to the next one. Therefore, the

structure of the processing chain is a result of the way in which the individual algorithms operate.

The literature review of § 2.4.5 showed that a number of algorithms already exist for extracting patches from pairs of images. However, the usefulness of these algorithms when applied to pairs of SAR and SPOT images is unclear. Therefore, they will be reviewed and tested with SAR and SPOT data so that methods can be defined for extracting similar patches from both images.

The literature review also showed that there are a number of techniques for matching patches. It is not clear if any of these techniques would produce reliable results when applied to patches extracted from SAR and SPOT images. Since matching techniques are far less well documented than extraction techniques, and to ensure accurate registration results, it is proposed that a new patch matching algorithm is developed, with emphasis placed on automation and reliability.

Edge extraction and edge matching algorithms have also been described in § 2.4.5. Edge extraction is well documented (Pratt, 1991; Gonzalez and Woods, 1992) and requires no further research in this study. A reliable method of edge matching has been described by Maitre and Wu (1989) and implemented successfully by many other researchers (Newton et al., 1994; Morgado and Dowman, 1997; Vohra et al., 1996). Therefore, research into the refinement of the registration using edge matching will be limited to the testing of the algorithms to determine the optimum method of implementation.

2.5.4 Conclusions

Automatic image registration is a huge subject which would benefit from research in almost any area. After considering an ideal system of automatic registration it was decided that this study would concentrate on the development of new, robust and reliable algorithms for extracting and matching features extracted from SAR and SPOT images. The development of the whole system, including decision making and validation algorithms, would make an interesting project for future work, but has not been included in this thesis. Consequently, the main aims of this study are:

1. to investigate and develop automatic feature extraction algorithms for use with SAR and SPOT data;
2. to develop an automatic patch matching algorithm for use with features extracted from SAR and SPOT data;

3. to integrate the patch extraction and patch matching procedures;
4. to investigate the possibilities of registration of both small and large images using the patch matching results; and,
5. to investigate the possibilities of extending the matching results to allow for full orthorectification of the data sets.

As a result, the layout of the remainder of this thesis is as follows: chapter 3 describes the acquisition of SAR and SPOT data; chapter 4 describes radiometric preprocessing of the images; chapter 5 investigates feature extraction; chapter 6 investigates patch matching; chapter 7 investigates edge matching; and chapter 8 applies the techniques developed to full scene images.

Chapter 3 DATA ACQUISITION AND IMAGE SELECTION

3.1 Introduction

This chapter describes in detail the systems that acquired the data used in this thesis: the SAR sensor onboard the ERS-1 satellite, and the optical linescanner on the SPOT platform. It is important to know how images are generated in order to be able to process them in an accurate and intelligent manner. For example, knowledge of image formation leads to an understanding of the geometric properties of the image, which is essential when applying geometric correction algorithms. Similarly, details of the imaging system can be used to determine the radiometric properties of the image, and hence increase interpretability. Without sensor information, it becomes much more difficult to understand the exact geometric and radiometric properties of an image, and therefore more difficult to model and correct the inherent inaccuracies. The importance of sensor modelling in image registration is discussed further in McGlone (1994).

3.2 Acquisition and properties of SAR data

3.2.1 Introduction

Radar remote sensing is a comparatively new remote sensing tool which comprises a number of different sensors and data products. Some of the microwave sensors are active, which mean they image the target by illuminating it with radiation, whilst others are passive, just measuring the extremely small amounts of microwave radiation emitted by the objects being imaged. Products derived from data acquired by radar sensors come in different forms, such as images, interferograms, or elevation models. In this project, the only relevant data product is image data acquired by an active microwave sensor, namely synthetic aperture radar, or SAR.

SAR images have been in use for many years. Early SAR sensors were mounted on aircraft, but since the launch of Seasat-1 in 1978 civilian spaceborne SAR data has been available. SAR is an active sensor, which means the sensor provides its own source of

energy, or illumination. This leads to a number of advantages over conventional optical remote sensing systems, the two main advantages being the ability of the sensor to acquire images at night, and the ability of the microwave radiation to penetrate through most types of atmospheric conditions. These factors mean that SAR data can be acquired at any time, leading to more continuous cover than can be provided with optical sensors. A further benefit of radar remote sensing is that the images are acquired in the microwave region of the electromagnetic spectrum, complementing images acquired in the optical and infrared regions, and thus giving the end user a wider range of wavebands to choose from when selecting imagery.

As mentioned above, the first civilian spaceborne SAR sensor was Seasat-1. This was followed by the experimental Soviet sensor Cosmos-1870 and the ‘commercial’ Soviet sensor Almaz-1. Data from Cosmos was not generally available, but Almaz provided eighteen months worth of data before returning to Earth. Soon after the launch of Almaz, the European Space Agency launched ERS-1 on the 17th July 1991, which operated successfully well beyond its intended lifetime of three years. In fact, so successful was ERS-1 that it was still operating normally when its successor ERS-2 was launched on the 21st April 1995. This gave ESA the opportunity to operate both systems in tandem, which proved enormously useful for sensor calibration and interferometric studies. The most advanced SAR sensor launched to date is on board the Canadian Radarsat platform. This sensor can operate in numerous different modes, giving different ground coverages and resolutions, and different polarities. It remains to be seen whether or not ESA’s next generation remote sensing platform, Envisat, will be able to better the sensors that have come before.

Although SAR images are of enormous value to the Earth observation community, their use is not as widespread for land application as optical images. This can be put down to a number of reasons, such as:

- ❑ low signal to noise ratio;
- ❑ lack of familiarity; and
- ❑ lack of continuity of acquisition.

Low signal to noise ratio

Due to the way in which the SAR sensor operates, described in more detail in the following sections, SAR image products are corrupted by speckle: a noise-like effect

that compromises the radiometric quality of the image and hence reduces the signal to noise ratio. This low signal to noise ratio means that objects are harder to resolve in a SAR image than in an optical image of similar pixel size. This immediately limits the possible uses of SAR images and users wishing to exploit remote sensing data will naturally choose types of imagery that give the best imaging potential; more often than not, this will not be SAR data.

Lack of familiarity

The acquisition and properties of SAR images are more complicated than those for optical images, and for the end-user perhaps a little bit further detached from reality. With optical data the user is familiar with the image product; firstly the image is similar to a photograph, with which all users would be familiar, and secondly it is generally possible to recognise features as if the observer were on board the platform and looking down at the Earth. This is not the case with SAR images, which do not have the same appearance as photographs, and also perhaps require a further quantum leap of the imagination to envisage how the sensor actually "sees" the ground. This lack of familiarity with SAR data has certainly led to people choosing to use optical rather than SAR data.

Lack of continuity of acquisition

Seasat was launched on 27th June 1978, but failed after only 100 days in orbit. Although it returned in excess of 100 million square kilometres of imagery, it was to be 9 years before the next civilian SAR ~~Satellite~~ (Cosmos-1870) was put into orbit. However, during this same period of time, optical remote sensing made huge advances and became the well-established tool of Earth observation. Perhaps if there had been a more continuous supply of spaceborne SAR data during this important period, SAR imagery would now be far more ubiquitous than is actually the case.

Due to the way in which the SAR sensor operates, SAR images have a number of interesting geometric and radiometric properties. An understanding of these properties should lead to the development of better processing algorithms, such as feature extraction algorithms or geometric correction algorithms. The next two sections describe the principal geometric and radiometric properties of typical SAR images.

3.2.2 Resolution

Since the technology of the SAR sensor was originally developed from its predecessor, the side looking airborne radar (SLAR) it seems appropriate to give a description of SLAR first as a prelude to the description of SAR.

The SLAR system is relatively straightforward: it consists of an antenna mounted on an aircraft, which transmits short pulses (of the order of 10^{-6} seconds) of microwave radiation and measures the time and the strength of the reflected signal (known as the backscatter). Each backscattered pulse of radiation is used to build up a line in the image; successive lines are built up from successive pulses of radiation and the forward motion of the aircraft (figure 3.1).

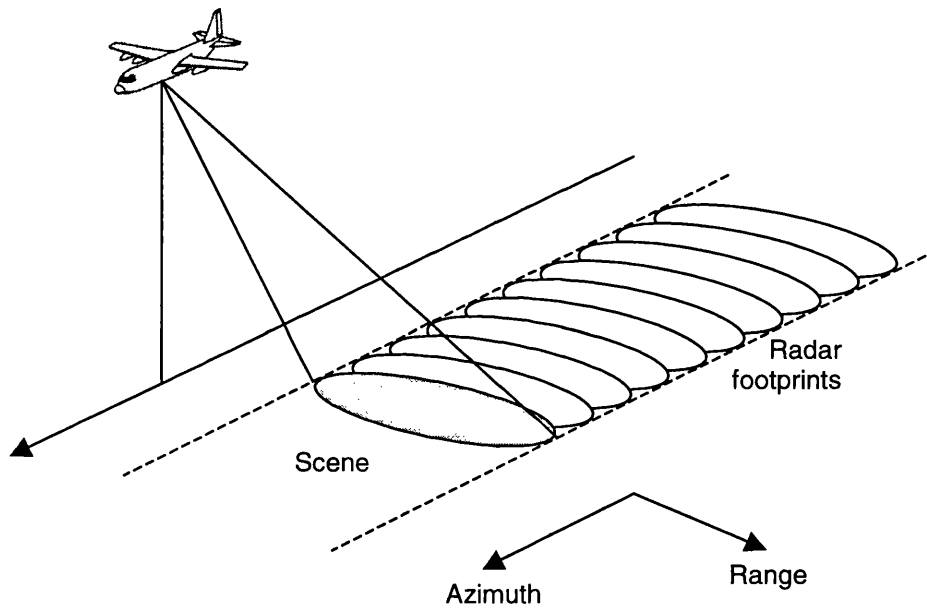


Figure 3.1 The SLAR sensor

Each individual backscattered pulse of radiation is made up of a number of echoes from different objects on the ground in the line of sight of the antenna. The resolving power of the sensor in the cross-track direction is determined by the ability of the sensor to separate these echoes. This is known as the range resolution. For two objects on the ground to be imaged separately and hence resolved, the pulse of radiation backscattered from each object must be received at the antenna separately. Therefore the closest the objects can be to each other, and still be resolvable is the distance that corresponds to half of the length of the pulse of radiation, τ , in the slant range. The distance corresponding to τ in the slant range (the slant range resolution) is given by

$$\text{Slant range resolution} = \frac{c\tau}{2} \quad (3.1)$$

where c is the speed of light. In terms of ground range, this is dependent on the look angle of the sensor (figure 3.2). Therefore, ground range resolution can be expressed as

$$\text{Ground range resolution, } R_r = \frac{c\tau}{2\sin\theta} \quad (3.2)$$

where θ is the look angle of the sensor (defined in figure 3.2).

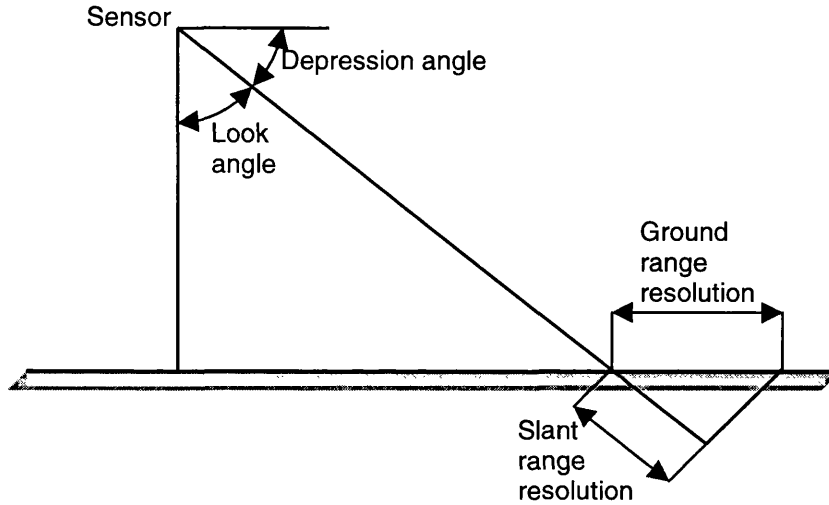


Figure 3.2 Geometry of the SLAR sensor

Since the pulse length is inversely proportional to the bandwidth of the *signal*, the range resolution of any pulsed radar system is fundamentally limited by the bandwidth of the transmitted pulse. A wide bandwidth can be achieved by using a short duration pulse, but the shorter the pulse, the lower the transmitted energy. This leads to a poorer the signal to noise ratio, hence worse radiometric resolution. Therefore, when designing a pulsed imaging radar system, a balance has to be achieved between geometric resolution and radiometric resolution.

The azimuth resolution of the SLAR sensor (measured in radians) is dependant on the beamwidth (not to be confused with bandwidth) of the antenna, and the ground range (figure 3.3).

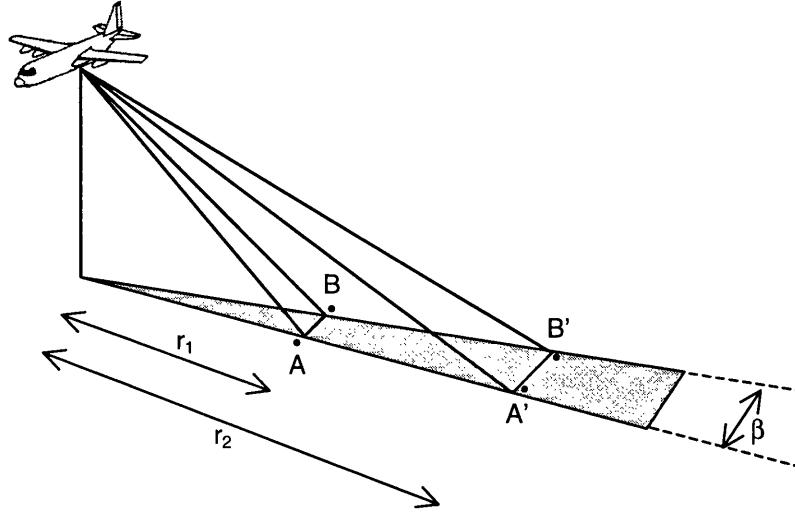


Figure 3.3 Azimuth resolution of the SLAR sensor (after Lillesand and Kiefer, 1994)

As the distance from the sensor increases, the beam of radiation fans out. At a distance r_1 from the sensor, the two objects A and B can be resolved since their separation is greater than the beamwidth. However, at a distance r_2 from the sensor, the two objects A' and B' (which have the same separation as A and B) cannot be resolved since they cannot be imaged separately: the beamwidth is now wider than their separation. Thus, azimuth resolution can be expressed as:

$$R_a = r\beta \quad (3.3)$$

where r is the ground range and β is the beamwidth of the antenna. The beamwidth is proportional to the wavelength of the radiation, λ , and inversely proportional to the antenna length, l ,

$$\beta = \frac{\lambda}{l} \quad (3.4)$$

Therefore, azimuth resolution can be expressed as

$$R_a = \frac{r\lambda}{l} \quad (3.5)$$

To improve the azimuth resolution of the SLAR sensor for a given wavelength, the length of the antenna has to be increased. This can be done in two ways: (1) by increasing the physical length of the antenna, or (2) by synthesizing the length of the antenna. The first method is obviously limited by the physical size of the platform, so the second method is used to improve the azimuth resolution, i.e. a synthetic antenna, or aperture, is created.

Although the SAR sensor only has a short physical antenna, it uses the forward motion of the platform to synthesize an antenna which is effectively very much longer. The principle depends on the use of coherent radiation, together with a precise knowledge of the transmit and receive point of the radar pulse. As the platform moves forward in its orbit, the distance from the antenna to the target (the slant range) changes continuously, and therefore the phase of the reflected signal changes as well. Since the phase is dependent on the geometry of the system, it is possible to correctly phase the return signals with respect to each other so that the effect is equivalent to them all having been received simultaneously by an antenna of length equal to the path length over which the radar signals were collected. In this way, the synthesized antenna can be thought of as a number of independently radiating antennae.

The azimuth resolution of a SAR sensor is determined by the Doppler bandwidth of the backscattered signal (for an explanation of Doppler shifting see Einstein, 1905.) Regions on the ground in front of the sensor will be approaching the sensor, and hence backscattered radiation will be Doppler shifted to a higher frequency. Similarly, radiation backscattered from objects on the ground behind the sensor will be Doppler shifted to a lower frequency, and radiation backscattered from objects in the line of sight of the sensor will have no Doppler shift (figure 3.4). By processing the backscattered radiation according to its Doppler shift, it is possible to synthesize a very small beamwidth.

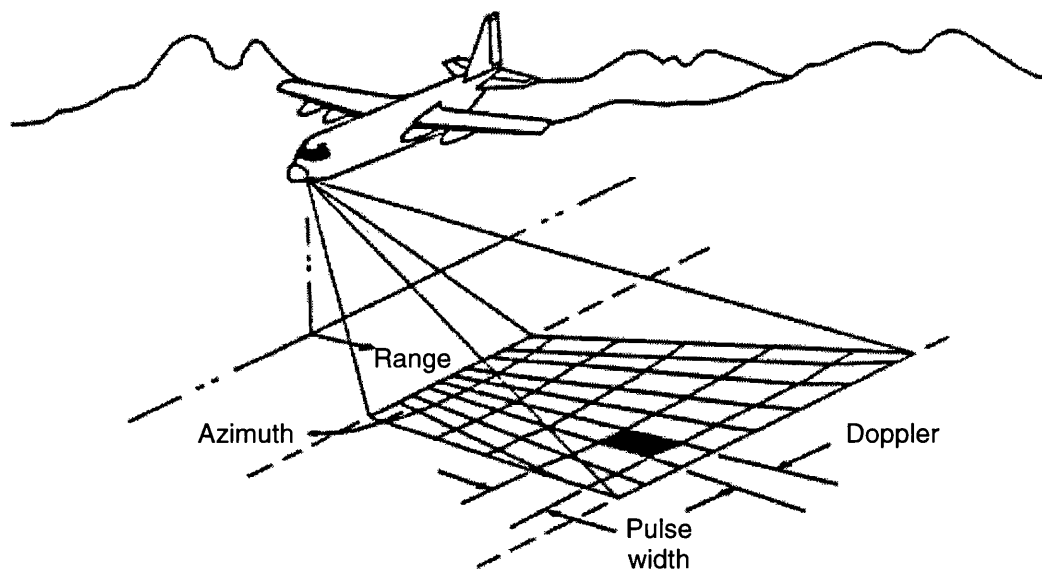


Figure 3.4 SAR and beamwidth

The consequence is that the azimuth resolution of a SAR sensor becomes a function of antenna length only:

$$R_a = \frac{l}{2} \quad (3.6)$$

Range resolution of the SAR sensor is the same as that of the SLAR sensor (given in equation 3.2).

3.2.3 Geometric distortions

The principal geometric effect in typical ground range SAR images is the displacement of objects due to relief (Leberl, 1990). Chapter 2 gave a brief description of how terrain affects the positions of objects in aerial photographs, and explained how distortions due to relief increase with increasing distance from the principle point. In SAR images, the opposite is true (figure 3.5). The projection of the building in the photograph causes the top of the building, A, to be imaged at *a*, further from the principal point than the base of the building, which is imaged at *b*. However, in the radar image, the top of the building is imaged at *a'*, which is closer to the nadir point than the base of the building, which is imaged at *b'*. This effect causes vertical objects to *layover* in the image, and is worse for objects close to the nadir of the sensor.

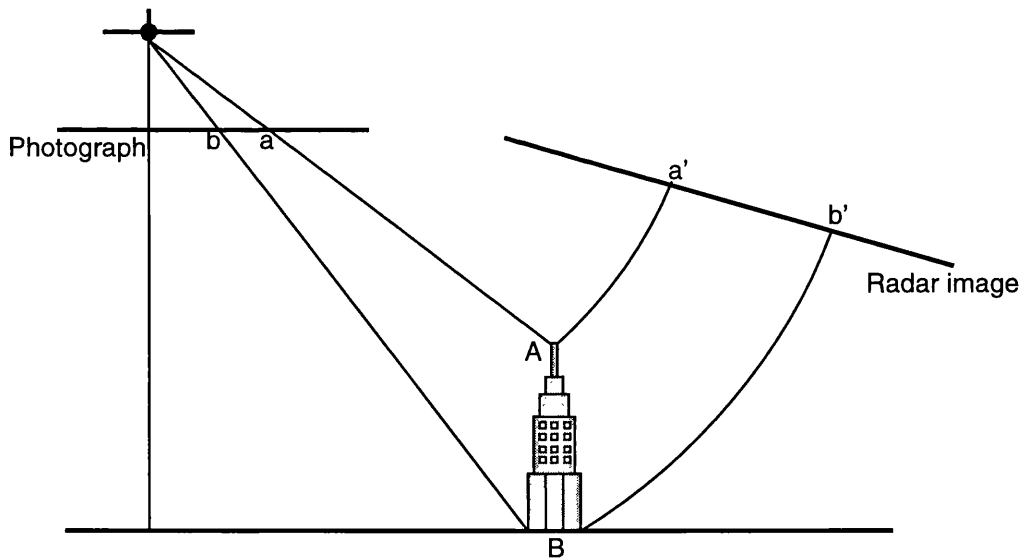


Figure 3.5 Distortions due to relief in optical and radar images
(after Lillesand and Kiefer, 1994)

3.2.4 Radiometric attributes

Speckle in SAR images originates from the coherent nature of the electromagnetic radiation that is used to form the image. A single resolution cell in the image, which may represent a region of 10m×10m on the ground, contains multiple scatterers. When the electromagnetic radiation interacts with these scatterers they produce echoes which may interfere constructively or destructively, depending on the exact structural content of the cell and the location of the antenna. The result is an image that appears to be

corrupted by random noise. In order to better understand how speckle corrupts SAR images, it is useful to create a statistical model. The model used is that of multiplicative noise, since this reflects the occurrence of speckle in SAR images. For multiplicative noise,

$$z_{ij} = x_{ij}v_{ij} \quad (3.7)$$

where z_{ij} is the observed value of pixel (i,j), x_{ij} represents the signal and v_{ij} represents the noise, with a mean of 1 and variance of σ_v^2 . The argument proposed by Lee (1981) showed that the standard deviation of the noise can be expressed in terms of the mean and variance of the observed pixel value:

$$\sigma_v = \frac{\sqrt{V_z}}{\bar{x}} \quad (3.8)$$

This result is extremely useful, since it gives a method of estimating the standard deviation of the speckle from the image itself; no other information regarding how the image was formed is required. By simply measuring the mean and standard deviation of a large number of pixels in a homogeneous region of any SAR image (in which all the pixel values are uncorrelated), it is possible to estimate the standard deviation of the speckle. It is important that the region is homogeneous since any features in the image will corrupt the results.

Although the standard deviation of the speckle can be estimated using equation 3.8, it can also be determined exactly if specific details of the image formation process are known. When the image is formed, multiple sub-images (or looks) can be combined to produce a multiple look image. This is done to reduce speckle, but in doing so, spatial resolution is also reduced. Therefore, spatial resolution is traded off against radiometric resolution. Lee (1986) demonstrated how the standard deviation of the speckle is dependent on the number of looks in the image, N , and how those looks are combined. When generating a multiple look image, either the amplitude of the range lines can be generated and averaged, or the intensities can be averaged before generating the amplitude. The two methods lead to a slightly different dependence of σ_v on N . Table 3.1 shows how σ_v varies with N .

Number of looks	σ_v (Processed by amplitude averaging)	σ_v (Processed by intensity averaging then computing amplitude)
1	0.5227	0.5227
2	0.3696	0.3630
3	0.3017	0.2941
4	0.2614	0.2536
6	0.2134	0.2061
8	0.1848	0.1781

Table 3.1 Dependence of σ_v on the number of looks, N (Lee, 1986).

The dependence of speckle on the number of looks was investigated using an ERS-1 PRI (precision product) SAR image of Southern France. The image was displayed so that a large homogeneous region could be found. The mean and standard deviation of the pixels were then measured in a large number of sub-regions (11×11 squares) within the main region. Values of standard deviation were grouped according to DN value, and the mean for each group was determined. These values were plotted against the DN value for the corresponding group (figure 3.6). A straight line was fitted to the data using a least squares method. The gradient of this line represents the standard deviation of the speckle.

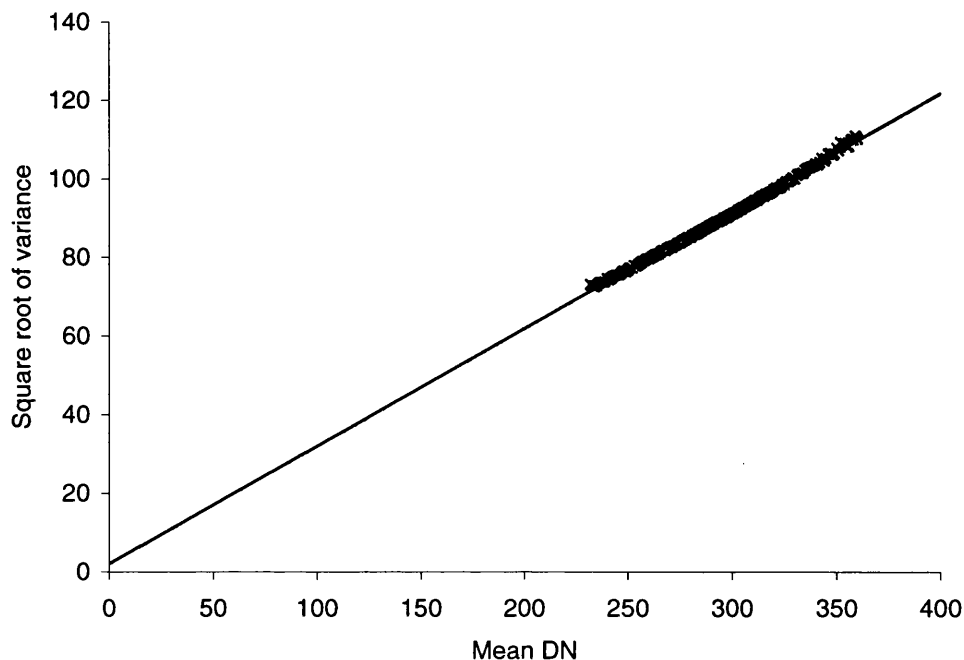


Figure 3.6 Mean and standard deviation of pixels in ERS-1 PRI SAR image

The standard deviation of the speckle was found to be 0.299, which would imply (from table 3.1) that the image is made up of 3 looks. This corresponds well to the true value which, according to ESA specifications for PRI image products, is 3 looks (ESA, 1993).

The importance of the standard deviation of speckle in SAR images becomes clear in the next chapter, which discusses speckle reduction. When designing algorithms to reduce speckle it is often useful to incorporate the statistical model given above. In practice this means knowledge of the standard deviation of the speckle is required in advance, in order to process the image.

Other obvious radiometric effects present in SAR images are shadow and sidelighting, caused by the way in which the microwave radiation interacts with the ground. Both of these effects are illustrated in figure 3.7.

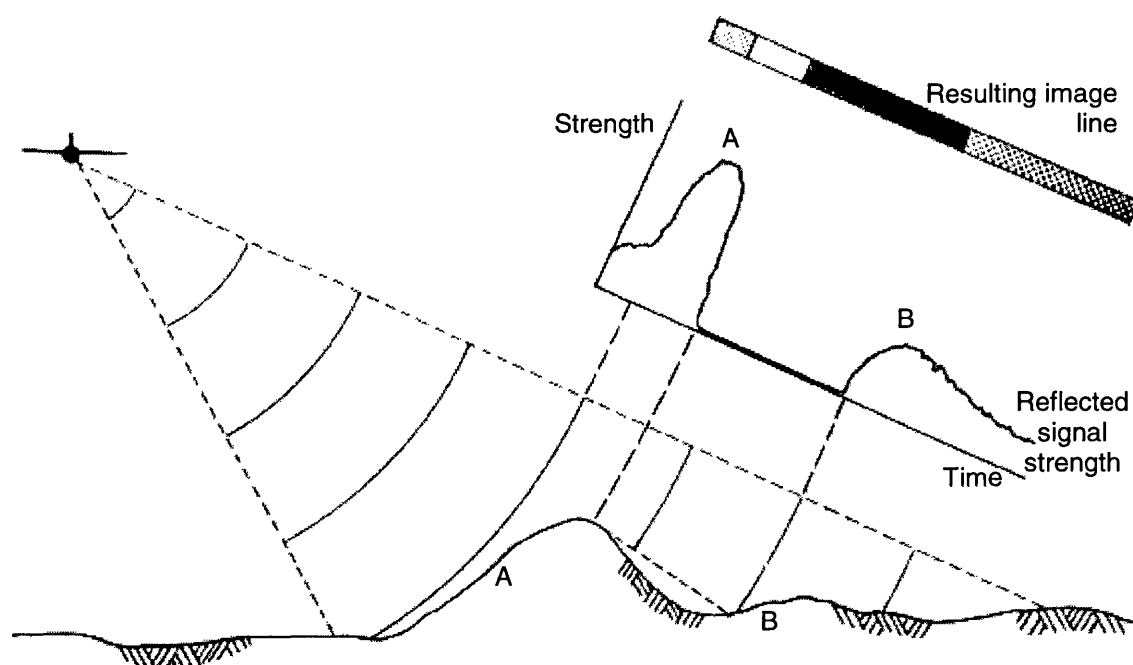


Figure 3.7 Sidelighting and shadow in SAR images (Lillesand and Kiefer, 1994)

As the above diagram shows, when the slope of the terrain facing the sensor is almost equal to the sensor look angle (i.e. the local incidence angle is small), the backscatter is high (slope A in figure 3.7). As the local incidence angle increases, the backscatter decreases (slope B). The second radiometric effect shown in figure 3.7 is shadow. The peak of slope A casts a radar shadow over the back of the slope, which is therefore not imaged; instead a region of black pixels appear on the image.

The radiometric properties of a SAR image are determined by how the microwave radiation physically interacts with the target. This interaction depends upon the dielectric properties of the target, the geometry of the target, and the frequency and

polarization of the incident radiation. However, this project does not consider multi-frequency or multi-polarization SAR data. Therefore, it is beyond the scope of this piece of work to discuss all the possible radiometric effects that may appear in an image, except to say that different wavelengths and polarizations interact differently with different targets, allowing different features to be imaged.

3.2.5 The ERS family of satellites

The European Space Agency's Environment Remote Sensing platform ERS-1 (shown in figure 3.8) was the first in a series of satellites dedicated to spaceborne environmental monitoring, the successors being ERS-2, and the hugely optimistic Envisat platform.

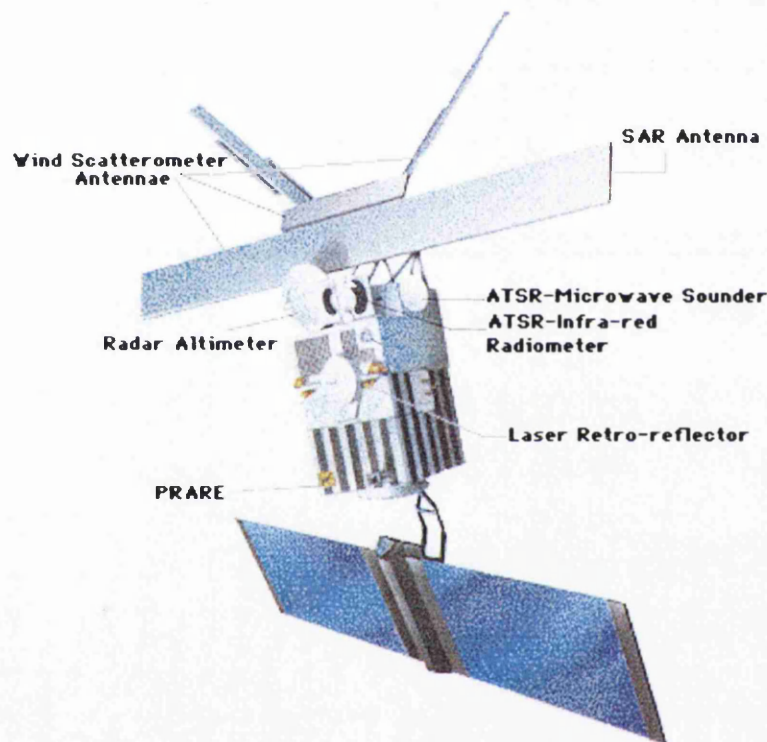


Figure 3.8 The ERS-1 satellite (© ESA)

The main instruments onboard the ERS platform are the active and passive microwave sensors, a radar altimeter, and an infrared radiometer. The active microwave instrument (AMI) has two modes of operation: image and wave. SAR images are acquired in image mode, in which the sensor has a spatial resolution of 26 m in range (across track) and between 6 and 30 m in azimuth (along track). Image data is acquired for a maximum duration of approximately ten minutes per orbit since the data rate is too high for onboard storage. It is therefore only acquired within the reception zone of a ground receiving station. The characteristics of the AMI and the orbital parameters of the platform are given in table 3.2.

Orbital parameters		AMI characteristics	
Type	Near-circular, polar, sun-synchronous	Frequency	5.3 GHz (C band)
Altitude	782 to 785 km	Bandwidth	15.55 Mhz
Descending equatorial crossing time	10:30am local time	Pulse duration	37.1 μ s
Inclination	98.52°	Polarization	Linear VV
Period	Approximately 100 minutes	Spatial resolution	6 to 30m along track; 26m across-track
Orbits per day	14.3	Swath width	100 km
Repeat cycle	3-day, 35-day and 176-day	Swath standoff	250 km to the right of the satellite track
Total mass	2157.4 Kg	Incidence angle	near swath 20.1° mid swath 23° far swath 25.9°

Table 3.2 Characteristics of ERS orbit and AMI specifications

The SAR antenna is aligned along the satellite's line of flight, directing a narrow beam sideways and downwards onto the Earth's surface (figure 3.9) to obtain strips of high resolution imagery of about 100 km in width.

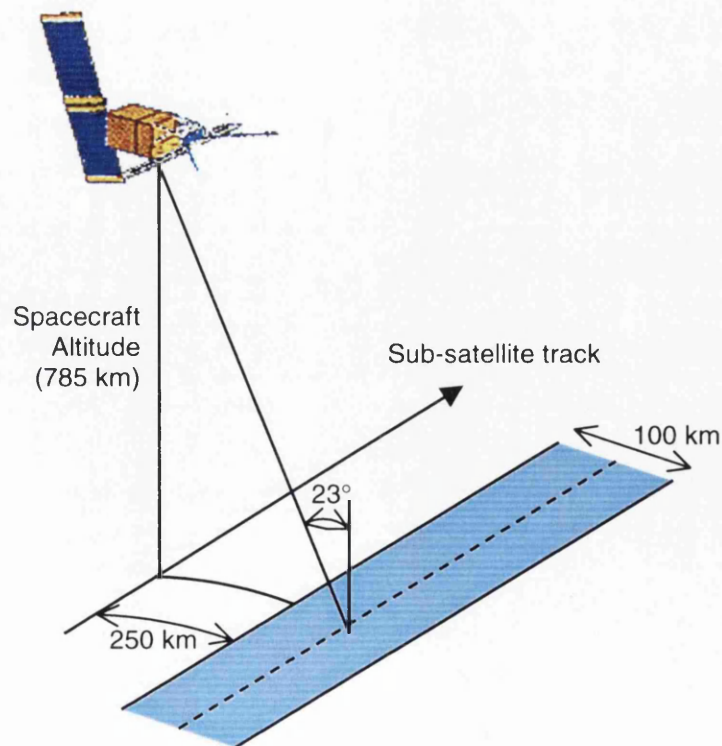


Figure 3.9 ERS orbital geometry (© ESA)

ERS-2 is identical to ERS-1 in almost every way. It was designed and built in order to provide continuity of data after ERS-1 had reached the end of its useful life. However,

ERS-1 proved to be unexpectedly successful and operated for much longer than anyone had predicted. This meant that ESA was able to operate both ERS-1 and 2 in tandem for some time, providing some very useful interferometric data. However, eventually at the end of May 1996, ERS-1 ceased regular operations, nearly five years after its launch. ERS-2 is still operating effectively and supplying useful data from all its instruments.

The next satellite in the ERS series, Envisat, will be the biggest satellite ever launched by ESA – with dimensions of 26m×10m×5m, it is literally huge (figure 3.10). Its payload is very similar to those of the ERS satellites, consisting of active and passive microwave instruments as well as an infrared radiometer, in order to ensure continuity of data. However, these instruments have been greatly improved and should therefore lead to far more useful data products. The sensor of most interest here is the AMI, known as ASAR (advanced synthetic aperture radar).

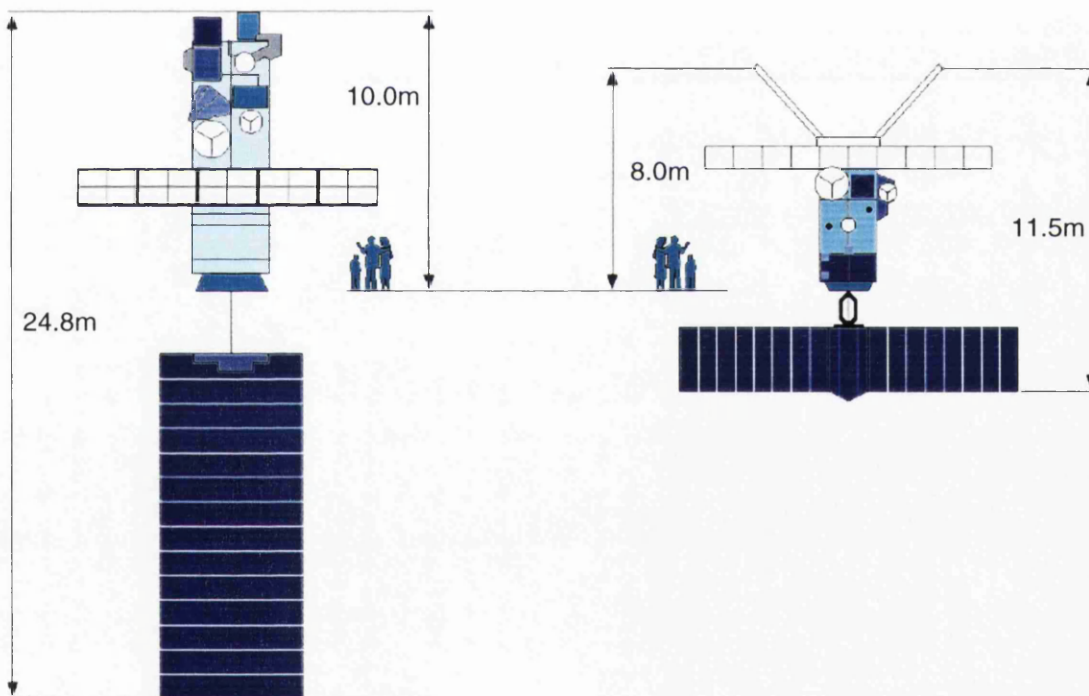


Figure 3.10 The Envisat platform (left) and ERS platform (right) (© ESA)

The ASAR instrument has five different operating modes, described in table 3.3, that can be used for imaging the Earth, depending on the application. This makes the ASAR considerably more adaptable than the SAR sensor of ERS-1, but still providing continuity of data.

	Image	Wide swath	Alternating	Wave	Global monitoring
Polarization	VV, HH	VV, HH	VV, HH, VH, HV	VV, HH	VV, HH
Spatial resolution	30m across 30m along	150m across 150m along	30m across 30m along	10m across 10m along	1km across 1km along
No. of looks	~4	~12	~2	1	~7
Swath width	Up to 100km	Up to 400km	Up to 100km	5km vignette	Up to 400km

Table 3.3 Specifications of Envisat's ASAR instrument

3.3 Acquisition and properties of SPOT data

3.3.1 Introduction

All of the optical data used in this study are SPOT panchromatic images acquired by the HRV (high resolution visible) sensor on the SPOT satellite. Since this sensor is a pushbroom scanner the first part of this section describes the principles of pushbroom scanners, whilst later sections describe the details of the SPOT satellite series, and the geometric and radiometric attributes of SPOT images.

3.3.2 The pushbroom sensor

A pushbroom (or along-track) scanner uses an linear array to build up an image of the ground line by line. As the platform moves forward the optics of the sensor project the image of a strip of ground onto the linear array. The array works by recording the intensity of this image line at discrete points along the line using CCD (charge coupled device) elements. The output of each element, which represents a pixel in the image, is stored electronically. When the sensor has moved forward a sufficient distance the next strip of ground is projected onto the array, and this in turn is recorded. By using the forward motion of the platform a two dimensional array of data (in other words, an image) is built up over a period of time (figure 3.11).

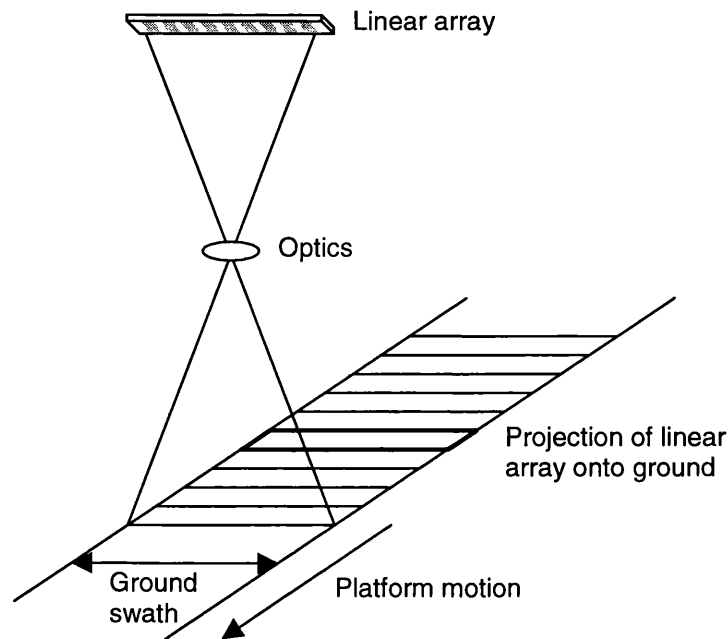


Figure 3.11 The pushbroom scanner

The linear array can contain over 10000 individual CCD elements. For imaging at more than one wavelength, multiple arrays are needed, one for each spectral band, each of which would be placed in the optical focal plane. A pushbroom scanner has a number of advantages over other electro-optical scanning devices. The linear array is a very stable device that requires little power to operate, has no moving parts and is relatively small and lightweight. Each of these attributes make it a very reasonable device to use on spaceborne platforms, where size, weight and longevity are key requirements. From a geometric perspective, the linear array has a fixed geometry which can be easily modelled. Radiometrically, it has the advantage that many different areas on the ground are imaged simultaneously, meaning that each CCD can spend longer imaging each individual region on the ground. This leads to greater energy being received by the sensor and hence a stronger signal and greater radiometric resolution. One disadvantage is that each CCD element has to be individually calibrated, but since this is not a significant problem, it does not compromise the usefulness of pushbroom sensors.

Pushbroom sensors have been employed on both airborne and spaceborne platforms. The best known spaceborne pushbroom scanner is the SPOT system, but the first to fly in space was the Modular Optoelectronic Multispectral Scanner (MOMS) built by Germany and launched on the Space Shuttle in 1983.

3.3.3 Image characteristics

Since optical linescanner images look similar to aerial photographs, it is easy to fall into the trap of assuming the geometric inaccuracies in these two data types are very similar.

In fact they are very different, the geometric distortions present in linescanner imagery being very much more severe than those in aerial photography. The reason for this is that linescanner imagery is acquired over a period of time, whereas an aerial photograph is a snapshot in time. The five major causes of geometric distortions are:

- ❑ off nadir viewing;
- ❑ Earth curvature;
- ❑ changes in flying height;
- ❑ Earth rotation; and
- ❑ changes in sensor attitude.

Distortions due to off nadir viewing exist because the sensor to target distance is different at nadir and off nadir. As the sensor to target distance increases (i.e. off nadir viewing) the line sampling interval increases. This causes the pixel size (in ground coordinates) to increase. For SPOT, the pixel size at nadir is 10m, whereas off nadir it is 13.5m. Since SPOT is able to view off nadir by as much as 27° distortions due to off nadir viewing are especially significant. Earth curvature also contributes to the scale distortions in the across track axis. Changes in flying height distort the image since line acquisition time is a function of flying height; height variations cause scale variations in the along track axis. As the image is acquired, the Earth rotates, the effect of which is that successive scan lines are shifted to the west. The amount of shifting is a function of latitude, and is worst at the equator. Finally, changes in attitude of the sensor while the image is being acquired leads to further distortions: rolling will cause lines to be shifted parallel to the pushbroom axis, pitching will cause successive lines to be interchanged, and yawing will cause scanlines to be rotated.

It is because of these distortions that a sensor model should be employed to geometrically correct optical linescanner imagery accurately (see, for example, Gagan, 1987; Haan, 1991; Kratky, 1989; Westin, 1990; O'Neill, 1991). However, it is quite difficult to model all the distortions – changes in flying height and sensor attitude are difficult to assess for the SPOT satellite. This means that even when a sensor model is used, it is possible that there will still be geometric distortions present in the "geometrically corrected" image.

3.3.4 The SPOT satellite series

The SPOT system is a series of optical remote sensing satellites designed and built by the French national space agency (CNES) with input from Belgium and Sweden, with each satellite carrying two identical HRV sensors. The first SPOT satellite, SPOT 1, was launched in 1986 and operated for four years before it was removed from active service. SPOT 2, launched in 1990 is still operating, but SPOT 3, launched in 1993, failed after three years in orbit. SPOT 4 was launched in 1997 and appears to be functioning nominally. SPOT 5 is planned for launch in 2002. Due to increased demand for SPOT data, especially in the Northern hemisphere, both SPOT 1 and SPOT 2 were operational in 1997. Such is the success of the SPOT system that after ten years of operation approximately 5 million images have been archived.

Each satellite in the SPOT series has carried the same pair of sensors, the HRV instrument, except for SPOT 5 which will carry a much more advanced pair of HRV sensors. The HRV sensor is an along track linear scanning array which has two modes of operation: panchromatic and multispectral. Both sensors can operate in either mode, either simultaneously or individually. The presence of two identical sensors is one of the reasons that the SPOT system is unique. A summary of the HRV characteristics, as well as the platform orbital parameters, is given in table 3.4.

Orbital parameters		HRV characteristics	
Type	circular, sun-synchronous, phased orbit	Operating wavelength (panchromatic mode)	0.51 to 0.73 μm
Altitude	822 km	Operating wavelength (multispectral mode)	0.50 to 0.59 μm (green) 0.61 to 0.68 μm (red) 0.79 to 0.89 μm (NIR)
Inclination	98 deg	Pixel size (panchromatic)	10m (at nadir)
Revolutions per day	14 + 5/26	Pixel size (multispectral)	20m (at nadir)
Period	101 minutes	Viewing angle	$\pm 27^\circ$ from nadir
Cycle duration	26 days	Swath width	60 km (at nadir)

Table 3.4 SPOT orbital parameters and sensor characteristics

The second reason for the SPOT system being unique is the ability to manually point the sensor away from the nadir, giving the opportunity for oblique viewing. This is done

using a steerable mirror which is controlled from the ground and is moveable through a range of $\pm 27^\circ$. Figure 3.12 shows a cross-section through the HRV sensor. The advantages that a steerable sensor offers are discussed in the next section.

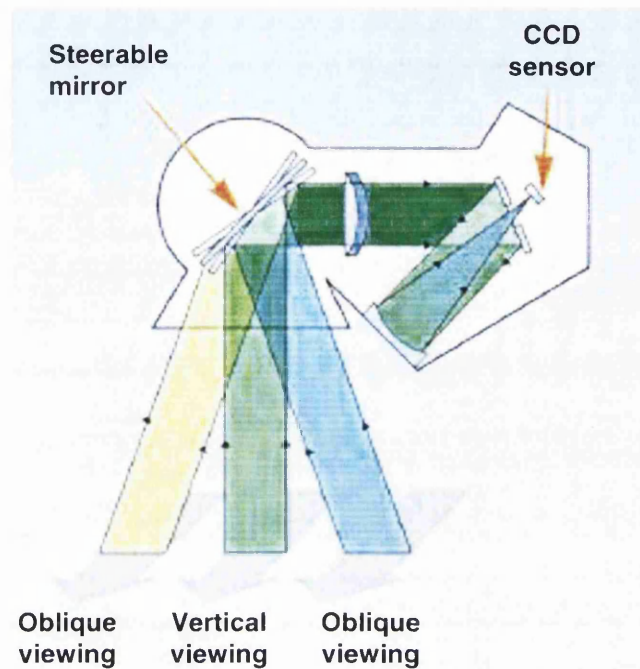


Figure 3.12 Cross section of the HRV sensor (© CNES)

In multispectral mode, SPOT 1, 2 and 3 have three spectral bands in the green, red and near infrared regions of the electromagnetic spectrum (SPOT 4 and 5 have more bands, but they are not described here) which were chosen to give the best possible imaging of vegetation and landcover. The panchromatic band covers most of the visible spectrum, and although useful for landcover monitoring, it is really designed for mapping purposes. The spectral bands are shown in figure 3.13.

Although the radiometric properties of the SPOT sensors are interesting, it is the geometric properties which really make this sensor stand out from the numerous other Earth observation sensors in operation.

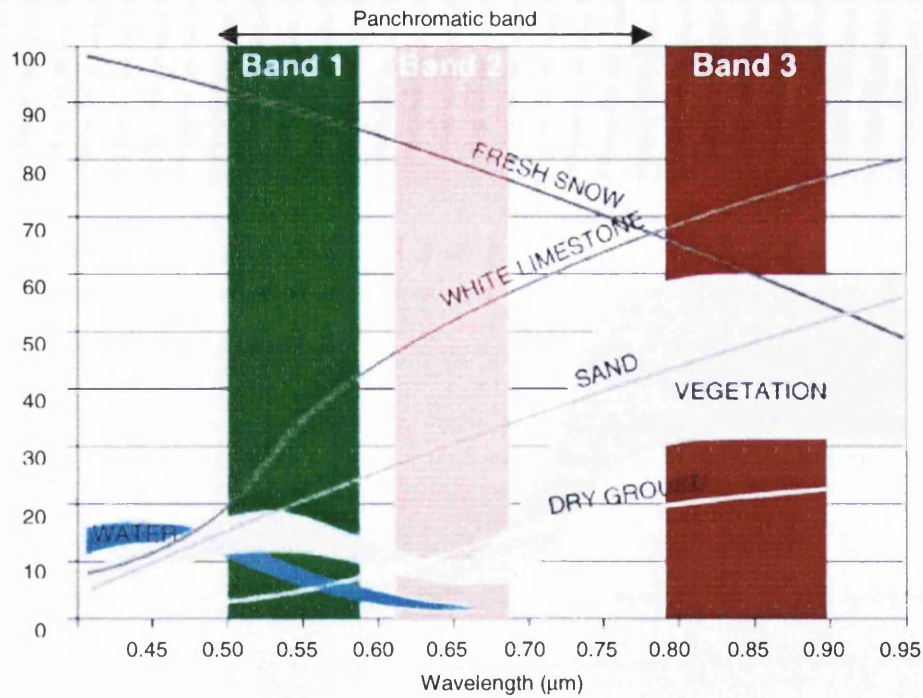


Figure 3.13 HRV spectral bands (© CNES)

The swath width of the HRV sensor at nadir is 60km. With both instruments pointing vertically downwards there is an overlap of 3km, giving a total coverage below the satellite of 117km (figure 3.14). Since this combined field-of-view is wider than the greatest distance between two adjacent tracks the sensors are able image every point of the Earth's surface during the satellite's 26 day cycle.

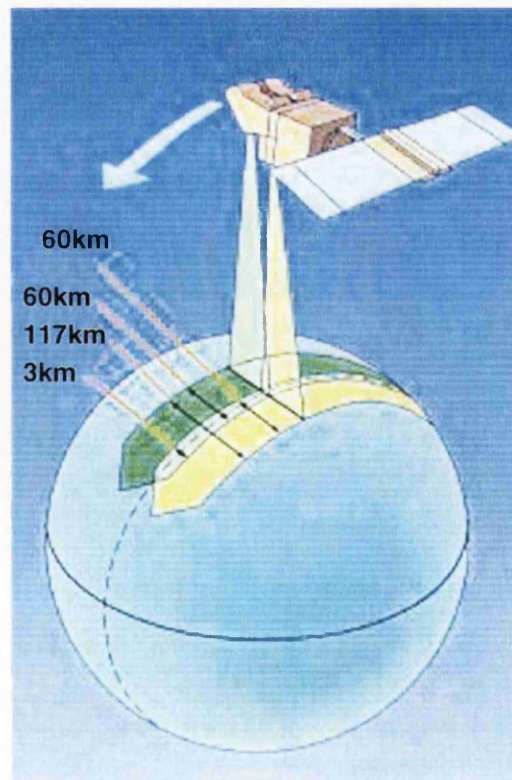


Figure 3.14 HRV viewing geometry (© CNES)

More important than regular complete coverage of the Earth is the opportunity of oblique viewing, which allows the capture of stereo-pairs of images (figure 3.15). This is done by imaging the same area on the ground from different orbits (i.e. different look angles), and thus creating a parallax difference between the images. For a pair of images acquired with viewing angles of 24° to the left and to the right of the sensor, a base to height ratio of 1.0 is achieved. An image acquired vertically below the satellite (viewing angle of 0°) and an image acquired at maximum viewing angle of 27° will give a base to height ratio of 0.5. As with aerial photography, these stereopairs can be used for topographic mapping and the generation of DEMs.

A second advantage of the oblique viewing capability is the increased ground coverage. SPOT is able to image any area on the ground within a 900 kilometre swath, leading to an increased viewing frequency for a given point during a given 26 day cycle. The repeat frequency varies with latitude, but at 45° the average time between passes is 2.4 days, with a maximum of 4 days and a minimum of 1 day.

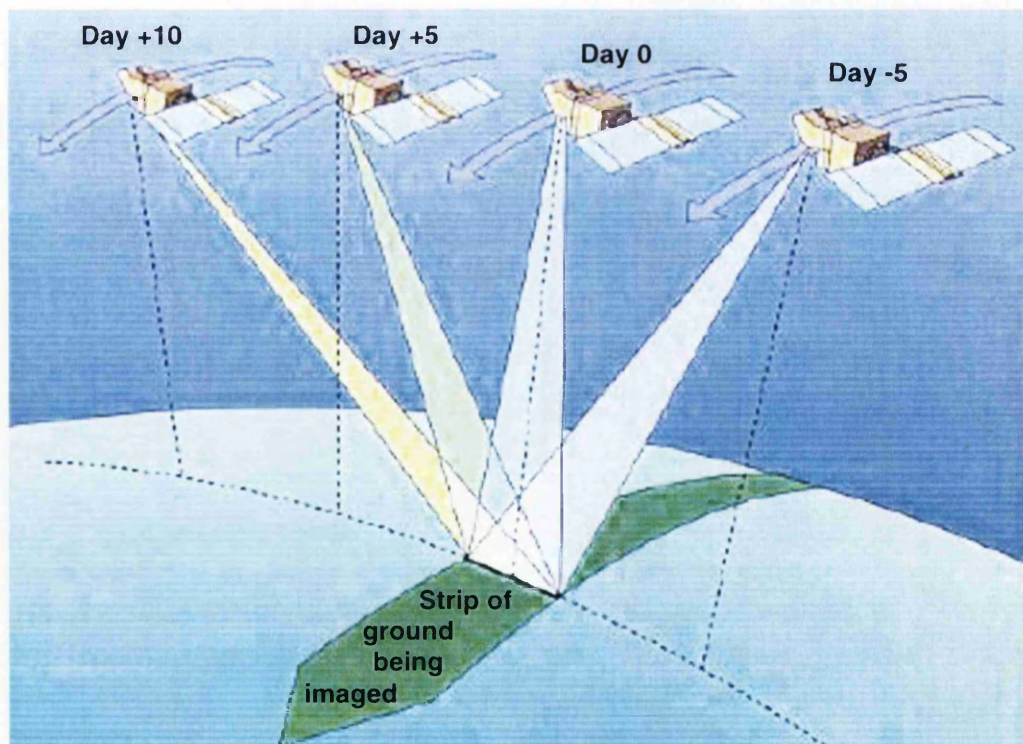


Figure 3.15 SPOT stereo viewing capability (© CNES)

Since SPOT 4 was launched in 1997 it has proved to be just as successful as the previous satellites in the SPOT series. It has a number of modifications, such as a new multispectral band in the mid-infrared ($1.58\text{--}1.75\text{ }\mu\text{m}$), onboard registration of all spectral bands (achieved by replacing the panchromatic band ($0.51\text{--}0.73\text{ }\mu\text{m}$) by band B2 ($0.61\text{--}0.68\text{ }\mu\text{m}$) operating in both a 10-m and 20-m resolution mode), and improved

recording capacity and telemetry capability. All the instruments have been validated, and apparently exceed their anticipated accuracies (SPOT Image, 1988).

SPOT 5, planned for launch in 2002, carries a new instrument: the high resolution geometry (HRG) sensor. Data will be acquired in four multispectral wavebands (B1: 0.50-0.59 μm ; B2: 0.61-0.68 μm ; B3: 0.79-0.89 μm ; and MIR: 1.58-1.75 μm) as well as the panchromatic waveband (0.51-0.73 μm) which was abandoned on SPOT 4. This will ensure continuity of the spectral bands established since SPOT 1. The HRG instrument offers much improved resolution in both the multispectral and panchromatic modes: 5m and 3m (instead of 10m) in panchromatic mode, 10m (instead of 20 m) in the three spectral bands in the visible and near infrared ranges, and 20m in the middle infrared band.

3.4 Image selection

3.4.1 Introduction

It was decided very early on in this study to use just two types of data: spaceborne SAR data, and SPOT data. The reason for doing this work is to investigate the possibilities of automatic image registration of two quite different types of data, of which SAR and SPOT are good examples. SAR data was chosen since it has a number of uses which are not always fully exploited, but could be when combined with optical data, whilst SPOT data was chosen since it is a principal data source for remote sensing. It would have been interesting to use images from other sensors, such as the Indian remote sensing satellite IRS-1C, or the Japanese ADEOS satellite, but in order to stay focused on the project the decision was taken to limit the number of data sets used. However, there is no reason why the techniques developed here cannot be applied to data from other similar sensors.

The study area used is Southern France, in the region of the Rhone Valley, shown in red in figure 3.16.



Figure 3.16 Study area in Southern France

This region was chosen for a number of reasons. Firstly, there is a large archive of data available from which images could be selected. Secondly, the area features many different types of landcover (mountains, agriculture, urbanization etc.). Thirdly, the area has been very well studied in the past by previous researchers. The combination of these reasons made this the obvious location of the test area.

3.4.2 SPOT data

The SPOT data used was acquired by SPOT 1 in 1986. The set of images cover areas of Southern France, from the Aix-en-Provence in the east to the Rhone Valley in the west. Table 3.5 gives details of the images. An example of one of the SPOT images, reduced in size, is shown in figure 3.17.

Acquisition date and time	Look angle and direction	Location of centre of image	Image size (columns/rows) (pixels)	Image size (x, y) (kilometres)
19/04/1987 11:01:01	25.8° Left	44°09'23" N 5°02'18" E	6000×6000	60.1×75.5
26/04/1987 10:26:18	24.8° Right	44°09'24" N 4°56'35" E	6000×6000	59.8×74.8
19/04/1987 11:01:09	25.8° Left	43°40'26" N 4°49'12" E	6000×6000	60.1×75.5
26/04/1987 10:26:26	24.8° Right	43°40'26" N 4°48'04" E	6000×6000	59.8×74.8

Table 3.5 Summary of SPOT images used in this study



Figure 3.17 Reduced size full scene SPOT image

3.4.3 SAR data

The SAR used in this study corresponds to the same area as the SPOT images described above. Details of the images are given in table 3.6, and an example of a reduced size image is shown in figure 3.18.

Acquisition date and time	Image product description	Location of centre of image	Image size (columns/rows) (pixels)	Image size (x, y) (kilometres)
30/12/1991 10:28:02	PRI	43.6N 4.9E	8000×8232	100×102.9
10/04/1992 22:01:13	RTM	44.0N 5.1E	8000×7767	100×97.9
06/05/1992 10:25:49	PRI	43.6N 5.5E	8000×8208	100×102.6
09/05/1992 21:49:52	PRI	43.9N 5.5E	8000×8203	100×102.5
16/11/1992 10:28:48	PRI	43.6N 4.8E	8000×8210	100×102.6

Table 3.6 Summary of SAR images used in this study

Although mostly ERS-1 SAR PRI (precision) images were used in this study, table 3.6 shows that one RTM (roll-tilt mode) image was used. The only difference between an RTM image and a PRI image is the sensor look angle during image acquisition. For most of the time, ERS-1 was operated with a look angle of approximately 23°

(producing amongst other products PRI images), but for a short period it was operated with a look angle of approximately 35° . Images acquired at this time were called RTM images. PRI and RTM images are almost the same, except that, for obvious reasons, the radiometric response of the terrain varies slightly. The difference between PRI and RTM images does not affect speckle reduction or feature extraction in this study in any way.



Figure 3.18 Example reduced size full scene SAR image (10:28:48)

3.4.4 Selected test sites

A number of small (but full resolution) square regions were cut from the SAR and SPOT images to be used as test sites for investigating the feature extracting and feature matching algorithms. They were chosen to reflect different landcover types and, more often than not, contained many strong features. These test images, shown in figures 3.19 to 3.22 below, are named after the regions on the ground that they depict. The SAR image is shown on the left and the SPOT image on the right. To facilitate the comparison of corresponding features, the SPOT images have been approximately aligned with the SAR images using a simple manual registration.



Figure 3.19 Beaucaire test images

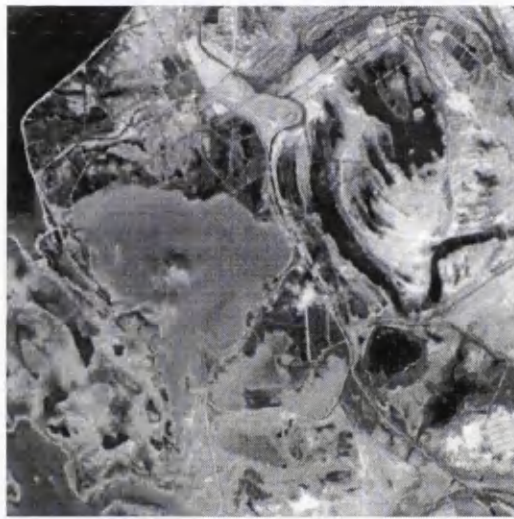
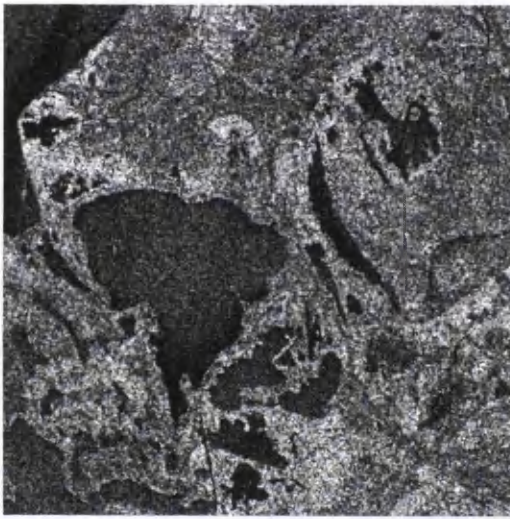


Figure 3.20 Camargue test images



Figure 3.21 Entressen test images

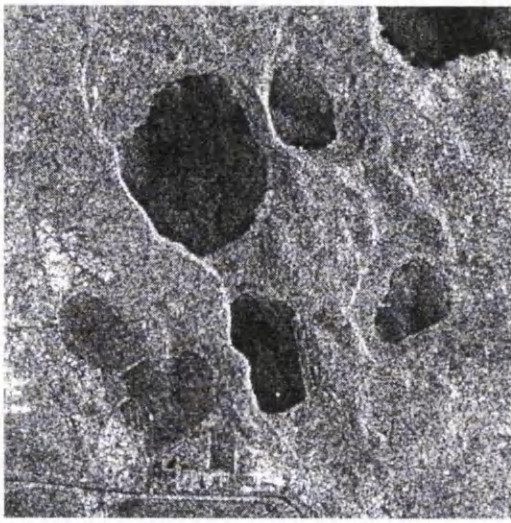


Figure 3.22 Istres test images

The Beaucaire image features a river running through the centre with a town on either side. Also clearly visible are linear features which are most likely railway lines (due to their strong response in the SAR image). The terrain is flat near the river, but more undulating further from the river. The Camargue image includes a number of lakes and mud flats near the coastline. Other than the lakes there are few other features. The terrain is generally very flat. The dominant features in the Entressen image are fields which are very well defined in the SPOT image, but much less clear in the SAR image. This region is also very flat. The last pair of images, the Istres test site, features lakes near the coastline, like the Camargue region, but in this scene the terrain varies quite considerably.

These four image pairs were selected to reflect different types of landcover and terrain. In addition they were chosen so that in each pair either the SAR or the SPOT image consisted of some strong features.

3.4.5 Conclusions

As the images in the above figures show, the region of Southern France around the Rhone Valley and Aix-en-Provence features numerous different types of landcover, many of which are clearly visible in either the SAR or SPOT images, but not necessarily both. It is the features that are visible in both that will be essential for the automatic registration of these data sets, underlining the importance of a wide range of landcover types in the test area.

There are other attributes of the data sets that have been taken for granted, but perhaps should be mentioned, such as acquisition dates, image resolution, and overlap. The SPOT images were acquired more than five years before the SAR data, which implies

that there could be some major differences between the images due to changes that took place on the ground between acquisition dates. It is not anticipated that this will be a problem in the matching of features since a large redundancy of features extracted means that there will be more than enough features to be matched. If any changes are detected, then this can be considered a very useful piece of information, and in itself justifies this work. The pixel size of both of the images are very similar to each other: 12.5m for SAR and 10m for SPOT, meaning that features visible in both images will be of a similar size (in image co-ordinates).

Regretfully there has been no opportunity of using data from other optical or radar sensors. It would have been interesting to compare results, but in order that this study should be realistic in its goals, a few images tested thoroughly is a more worthwhile target than a large number of images that have only been tested superficially.

3.5 Discussion

The purpose of this chapter was to introduce the subject of image acquisition in order to highlight some of the geometric and radiometric aspects of SAR and SPOT images. It is accepted that a knowledge of the formation of the images should lead to a better understanding of how to develop and apply image registration algorithms to specifically work with these images. However, the subject of image acquisition is literally huge, and it really has not been possible to cover more than a minute fraction of it here. A full description of imaging radar can be found in Henderson and Lewis (1998), but unfortunately there is no similar single reference which fully describes optical remote sensing.

The first section of this chapter showed how a knowledge of the formation of SAR images can lead to the development of rigorous processing techniques, both geometric and radiometric. An understanding of the imaging geometry explains well all the geometric distortions present in SAR images. In turn this has led to the development of techniques for removing these distortions and generating geometrically accurate image products (orthoimages). The subject of radar image geocoding was briefly introduced in the previous chapter, but a full description of the subject of radargrammetry can be found in Schrier (1993).

The second section of this chapter introduced optical linescanners in general, and the SPOT system in particular. The geometric distortions present in optical linescanner imagery can be simply explained by considering the geometry of the sensor and the

method of image acquisition, and removed from the image using a sensor model (Gugan, 1987). The main geometric distortions present are caused by the fact that the image is acquired over a period of time.

The radiometric properties of line scanner imagery are harder to quantify since they depend very much on which sensor is used to acquire the images. Different sensors operate in different parts of the electromagnetic spectrum, and each waveband has its own purpose: visible and near infrared wavebands for vegetation monitoring, middle infrared wavebands for geological mapping, and thermal infrared for oceanic and atmospheric applications. The purpose of this study is accurate image registration, which requires the highest resolution data, namely panchromatic data, which in the SPOT sensor straddles the visible wavelengths.

§ 3.4.4 introduced the SPOT satellite and the HRV sensor – the instrument which acquired the optical images used in this study. It can clearly be seen that this is an enormously powerful and successful remote sensing tool, especially in the field of topographic mapping. A number of studies have shown that topographic maps and DEMs generated from SPOT data are accurate enough for mapping at scales of up to 1:50 000 (Gugan and Dowman, 1988). SPOT is able to acquire images in a panchromatic waveband, and from different viewing locations, giving 10m resolution stereo pairs suitable for DEM generation and production of orthoimages. The method of geometric correction is exactly the same as the photogrammetric techniques used for aerial photography: the only supplementary data required is a set of ground control points. Although a DEM is also required, this can be generated from the SPOT data, as long as a stereopair of images is available. Techniques for automatic DEM generation from SPOT data are well established (see for example Day and Muller, 1989), and are now available as a part of many image processing software packages. It is for these reasons that SPOT data has been used in this study, rather than other optical sensors such as Landsat, which offer much better radiometric imaging, but far less advanced geometric viewing capabilities.

The final topic covered in this chapter was image selection. The study area was chosen carefully to ensure that a good selection of suitable images were available. One of the main limitations of image registration based on feature matching is that for it to be successful it is essential to have images which contain many different types of features. It goes without saying that featureless images will be impossible to match using feature

based image registration, and probably not possible using any other technique. The region in Southern France meets all the requirements in terms of visibility of features within images from both SAR and optical sensors.

Chapter 4 RADIOMETRIC PREPROCESSING

4.1 Introduction

Now that the images which are going to be used in this study have been introduced it is possible to begin describing the processing of these images. Before feature extraction algorithms can be applied to the raw data, the images have to be radiometrically preprocessed in order to improve the success rate of the feature extraction algorithms.

The presence of speckle in the SAR images means that SAR preprocessing algorithms (smoothing and speckle reduction algorithms) have to be able to remove a lot of ‘noise’ whilst preserving the features in the image. This is a very tall order, and for a number of years researchers have been trying to develop ideal speckle reduction algorithms. The discussion of speckle presented in the previous chapter explained how speckle could be mathematically modelled – this chapter explains how this modelling can be used to develop speckle reduction algorithms. However, it is not essential to model speckle in this way in order to develop speckle reduction algorithms and a number of non-rigorous, but very useful, algorithms are presented.

The preprocessing of SPOT images is very different from the preprocessing of SAR images since the signal to noise ratio of SPOT data is very much higher than that of SAR data. In fact, very little preprocessing is required, but a few algorithms are presented which enhance the features in the image and therefore should improve the success of feature extraction algorithms described later in this study.

4.2 Preprocessing of SAR data

4.2.1 Introduction

The usefulness of SAR images is clear, but the noise-like speckle which corrupts the images makes them difficult to interpret. To improve the interpretability, and therefore usefulness of SAR images, it is essential to try to remove the speckle. A very large number of speckle reduction algorithms (or filters) have been developed over the years, and this section aims to summarize some of the more important ones, as well as

introduce some new ideas on speckle reduction. For ease of discussion, the speckle reduction filters described here have been divided into four groups (although this division is not rigid):

- ❑ simple;
- ❑ adaptive, rigorously modelling the statistical properties of the speckle;
- ❑ adaptive, but not modelling the statistical properties of the speckle; and
- ❑ improved traditional filters.

4.2.2 Simple speckle reduction filters

Simple filters are those which perform exactly the same algebraic operation on all pixels in an image. They have the advantage that they are made up of basic algebraic operations, so are therefore easy to implement and quick to produce results. Their disadvantage is that they take no account whatsoever of the physical principles behind the formation of a SAR image, and therefore do not necessarily produce the best results.

The principal simple speckle reduction filter is the mean, or moving average, filter. The 3×3 mean filter is implemented by convolving the image with the kernel:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (4.1)$$

In practice, this is done by moving a 3×3 window across the image, pixel by pixel. The digital number (DN) of the pixel in the centre of the window is replaced by the mean of all nine pixel DNs in the window. This process is repeated for every pixel in the image. This filter can be thought of as a special case of the general 3×3 parametric low pass filter (Gonzalez and Woods, 1992), defined by the kernel:

$$\left(\frac{1}{b+2} \right)^2 \begin{bmatrix} 1 & b & 1 \\ b & b^2 & b \\ 1 & b & 1 \end{bmatrix} \quad (4.2)$$

Further variations of the mean filter can be generated by changing the size of the kernel (e.g. 5×5, 7×7 etc.) or changing the ‘shape’ of the kernel. The 3×3 plus-shaped mean filter is defined by the kernel:

$$\frac{1}{5} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4.3)$$

Two further variations of the mean filter are the threshold mean filter (Pratt, 1991) and the k -average filter (Davis and Rosenfeld, 1978). The threshold mean filter will only replace the central pixel with the mean of the eight surrounding pixels if the difference between the two is above some user defined threshold, else no change takes place. The k -average filter takes the mean of the k pixels nearest in value to the central pixel. For a 3×3 kernel, the value of k is usually set to 6 for best results.

Another very important and often used simple filter is the median filter. It works by moving an $n \times n$ window across the image, and replacing the central pixel with the median of the n^2 elements of the window. This filter proves to be very successful since spurious DNs in the window do not affect the output.

The last simple filter introduced here is the Gaussian filter. The image is smoothed by convolving it with a kernel generated from a 2 dimensional Gaussian function, $G(x,y)$, where

$$G(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4.4)$$

In practice, σ^2 is set to one third the size of the operator. For example, for an operator of size 5, the convolution kernel is:

$$\frac{1}{1071} \begin{bmatrix} 0 & 11 & 26 & 11 & 0 \\ 11 & 47 & 109 & 47 & 11 \\ 26 & 109 & 255 & 109 & 26 \\ 11 & 47 & 109 & 47 & 11 \\ 0 & 11 & 26 & 11 & 0 \end{bmatrix} \quad (4.5)$$

From the kernel, it can be seen that pixels close to the central pixel play more of a role in determining the output of the filter than pixels further away from the central one.

To compare the results of the filters described above, they have each been applied to a 200×200 pixel test image taken from the ERS-1 PRI SAR image of Beaucaire. The test image (shown in figure 3.19 in the previous chapter) features a portion of the Rhone

river near the town of Beaucaire. The river appears as a dark feature passing through the centre of the image. Urbanization (the very bright regions) can be seen on both sides of the river, and bridges can be seen crossing the river. It is difficult to clearly make out field patterns, but reference to a map of the region shows that the mid-grey regions north of the urbanization does in fact represent agriculture.

Figure 4.1 shows the Beaucaire test image after the mean filter has been applied. The image on the left has been generated using a 3×3 kernel, and the image on the right using a 9×9 kernel.

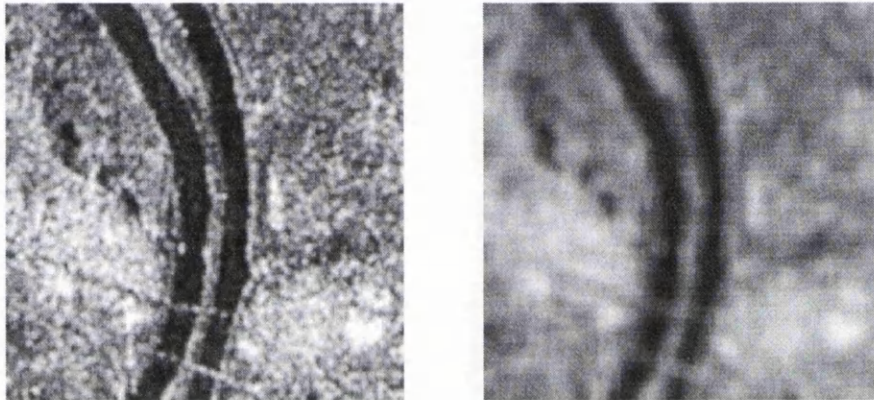


Figure 4.1 Mean filter applied to Beaucaire image (left: 3×3 , right: 9×9)

Figure 4.2 shows the Beaucaire test image after the plus shaped mean filter has been applied. As with the above example, the image on the left has been generated using a 3×3 kernel, and the image on the right using a 9×9 kernel.

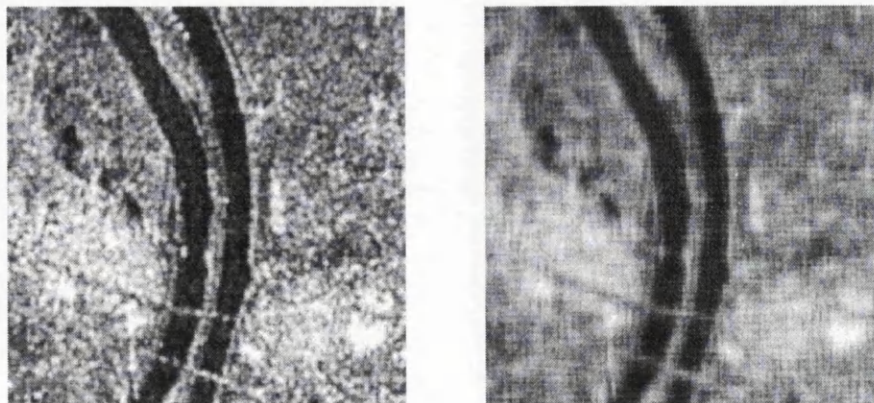


Figure 4.2 Plus shaped mean filter applied to Beaucaire image (left: 3×3 , right: 9×9)

In figure 4.3 the 3×3 and 9×9 median filter has been applied to Beaucaire test image.

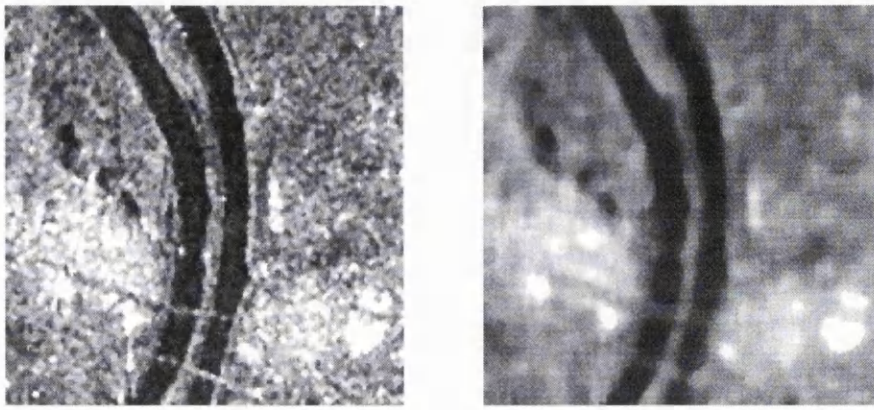


Figure 4.3 Median filter applied to Beaucaire test image (left: 3×3 , right: 9×9)

Lastly, figure 4.4 shows the Beaucaire test image after Gaussian smoothing using a 3×3 and 9×9 kernel.

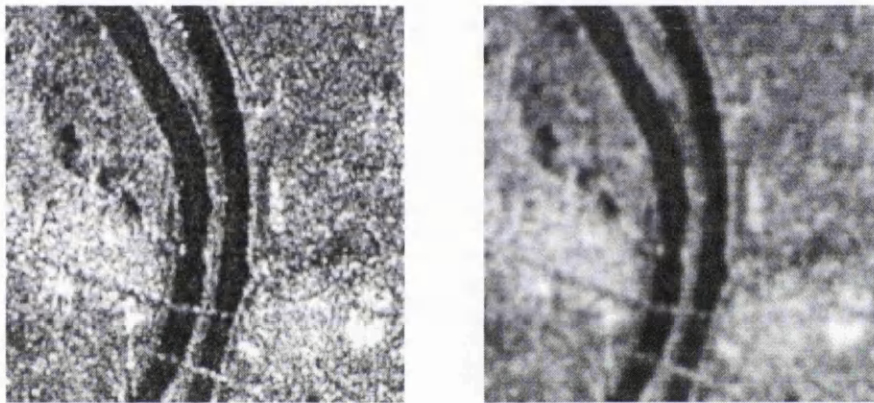


Figure 4.4 Gaussian smoothing of Beaucaire test image (left: 3×3 , right: 9×9)

By looking at the above smoothed images it is immediately obvious that there is much more smoothing if the filter is applied over a larger window, irrespective of which filter is being used. It is also clear that the mean filter gives the most smoothing, and the others all slightly less. However, the cost of good speckle reduction is the loss of detailed features in the image, or in the case of the 9×9 plus shaped mean filter, the addition of a new texture effect to the image which was not there before. The best compromise of the above filters is the 3×3 median filter, which smooths much of the speckle away but retains many of the detailed features.

The list of simple filters described here is by no means exhaustive. Over the years there have been many different types of simple filters developed, a complete summary of which would be too long to be included in this study.

4.2.3 Adaptive, rigorous filters

Adaptive filters, as their name suggests, adapt the way in which they process the image depending on local conditions. The four adaptive filters described here all use a rigorous

statistical model of the speckle. The statistical model given in § 3.2.4 assumes the noise intensity between pixels is statistically independent (i.e. the image is uncorrelated). Therefore, any filter relying on a multiplicative noise model should only be applied to an uncorrelated image. However, the ERS-1 PRI SAR images used throughout this work are correlated, so processing must be carried out to eliminate the correlation in the image. This can be done by averaging down the image as intensity data whilst preserving the statistical properties. When the despeckling has been completed, the image can be restored to its original size again. In the examples described below, this procedure was carried out using tools from the CAESAR image processing toolkit (Caesar, 1996).

The four adaptive filters described below are implemented by moving a fixed size window across the image. The central pixel is assigned a new value which is dependent on the image statistics within the window, as well as the global image statistics. In all the examples below, the following definitions apply:

x = original value of central pixel in window

\hat{x} = new value pixel assigned to central pixel in window

μ = mean of pixel values within window

σ_x = standard deviation of pixels within window

V_x = variance of pixels within window

σ_v = standard deviation of speckle noise

V_v = variance of speckle noise

The Frost filter (Frost et al., 1982) uses the following algorithm to process the image:

$$\hat{x} = mx \quad \text{and} \quad m = \frac{A\sigma_x}{\mu} \exp\left(\frac{-B\sigma_x}{\mu}\right) \quad (4.6)$$

where A is a normalization factor, and B is a constant to be set empirically. The Lee filter (Lee, 1981) processes the image by applying the following algorithm to each pixel in the image:

$$\hat{x} = \mu + \left(\frac{V_x}{V_x + \mu^2 V_v}\right)(x - \mu) \quad (4.7)$$

The Kuan filter (Kuan et al., 1985) is a more generalised version of the Lee filter, and is defined by the equation:

$$\hat{x} = \mu + \left(\frac{V_x}{V_x + V_v(\mu^2 + V_x)} \right) (x - \mu) \quad (4.8)$$

The only difference between the Lee algorithm and the Kuan algorithm is the additional term in the denominator, which does not appear in the Lee algorithm due to the linear approximation that was made for the non-linear multiplicative noise model. The last algorithm introduced here is the maximum a posteriori (MAP) filter, proposed by Kuan et al. (1987). An estimate of the original pixel value, x , is obtained by maximising the *a posteriori* probability density function,

$$P\langle x|z \rangle = \frac{P\langle z|x \rangle P\langle x \rangle}{P\langle z \rangle} \quad (4.9)$$

with respect to x (Kuan et al., 1987). If $P\langle z|x \rangle$ is chosen to be a gamma function and $P\langle x \rangle$ is chosen to be Gaussian, then the following expression can be derived:

$$\hat{x}^2(\hat{x} - \mu) + \frac{\sigma_x^2}{\sigma_v^2}(\hat{x} - x) = 0 \quad (4.10)$$

The real root of this cubic equation gives the value of the output pixel.

As with the simple filters, each of the above adaptive filters have been applied to the Beaucaire test image. In each case a 5×5 window was used to calculate local image statistics. Figure 4.5 shows the results of the Frost and Lee filters, whilst figure 4.6 shows the results of the Kuan and MAP filters.

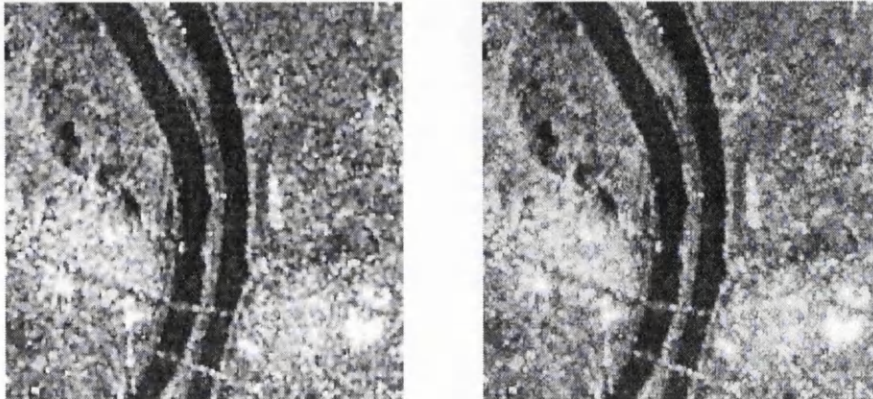


Figure 4.5 Frost filter (left) and Lee filter (right) applied to Beaucaire test image

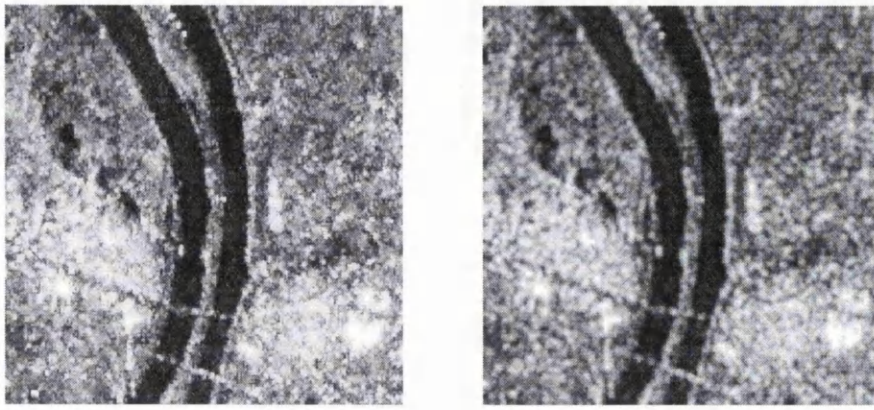


Figure 4.6 Kuan filter (left) and MAP filter (right) applied to Beaucaire test image

The results of these four filters are very different to the results generated by the simple filters in the previous section. The adaptive filters have retained much more detail than most of the simple filters, but have also removed less speckle. The Frost filter has retained a lot of 'spot' noise, which has been interpreted as being features rather than speckle. The Lee and Kuan filters have produced similar results, but the Kuan filter appears to give a slightly smoother result. The MAP filter has removed the most amount of speckle and retained the fewest features.

As with the simple filters, there are many more rigorous adaptive filters that have been developed over the years. To describe them all here would digress too far from the aims of this study, so only the principle ones have been introduced.

4.2.4 Adaptive, non-rigorous filters

The next group of speckle reduction filters to be described are the adaptive but non-rigorous filters. Only one filter in this group (the sigma filter) uses the statistical model of multiplicative noise to determine the standard deviation of the speckle, but at no stage is the speckle modelled - the standard deviation of the speckle is just used as a parameter in the processing. The three filters described in this section are the sigma filter (Lee, 1983), the most homogeneous neighbour (MHN) filter (Nagao and Matsuyama, 1979) and the modified k -average filter (Narasimha Rao et al., 1995).

The sigma filter operates by replacing the pixel being processed by an average of those neighbouring pixels whose DN values lie within two noise standard deviations of the original pixel. Hence, pixels which represent sharp spot noise are excluded from the summation. A second variation of this filter incorporates a threshold (k) which ensures that there are enough pixels in the averaging process from which to take a sensible mean. For a 5×5 mask, k is usually set to 2, so there must be more than two pixels from which to take the mean. For a 7×7 mask, k is usually set to 3.

The most homogeneous neighbourhood (MHN) filter works by taking a 5×5 window centred on the pixel being processed and generating nine separate masks over the window (figure 4.7). The mean and variance of each mask are determined and the DN value of the central pixel is replaced by the mean of the mask with the lowest variance. It is assumed that the lowest variance mask has the least chance of crossing an edge, so this algorithm should remove noise whilst preserving edges (hence the alternative name, edge preserve smoothing).

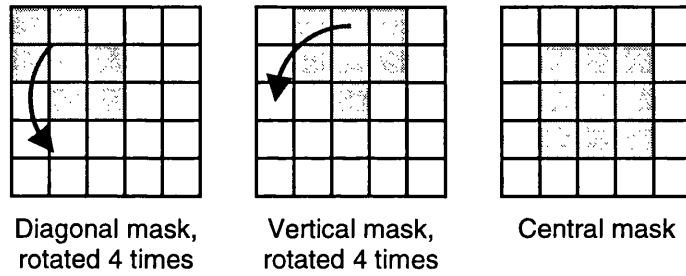


Figure 4.7 Nine masks used in MHN filter

The modified version of the k -average filter differs from the original version by allowing the value of k to change according to local conditions. In homogeneous regions a high value of k gives better speckle reduction whereas in regions of low homogeneity a low value preserves features better. The value of k is calculated using the expression:

$$k = n^2(1 - s) \quad (4.11)$$

where $s = (V_{loc} - V_{min}) / (V_{max} - V_{min})$, V_{loc} is the variance of local window, and V_{max} and V_{min} are the maximum and minimum variance of whole image, and n is the window size. However, the authors completely failed to realise that the variance across a SAR image can vary dramatically. A 512×512 sub-image was extracted from a full scene ERS-1 PRI SAR image, and the maximum and minimum variance, V_{max} and V_{min} were measured. It was found that $V_{min} \approx 60$, while $V_{max} \approx 8000000$; the mean of the variances was approximately 9000. Thus, applying the modified k -average filter using a 5×5 window to this image meant that for virtually all the pixels, the value of k was 25. There were literally only one or two pixels where k took a lower value. To solve this problem, the histogram of the image needs to be adjusted so pixels with very high DN values are eliminated, and therefore making the variances occupy a much smaller range. When this was done it was found that the values of k used in the smoothing of the image were spread across the range of 1 to 25 much more evenly. However, the smoothing now

depends very much on the extent to which the histogram is manipulated, making application of the filter very image-dependent, and hence comparison with other filters quite difficult.

The application of these three filters are shown below in figure 4.8. The modified k -average filter was applied to a version of the test scene with a modified histogram in order to generate a visual result, but there is no further analysis of this filter for the reasons given above.

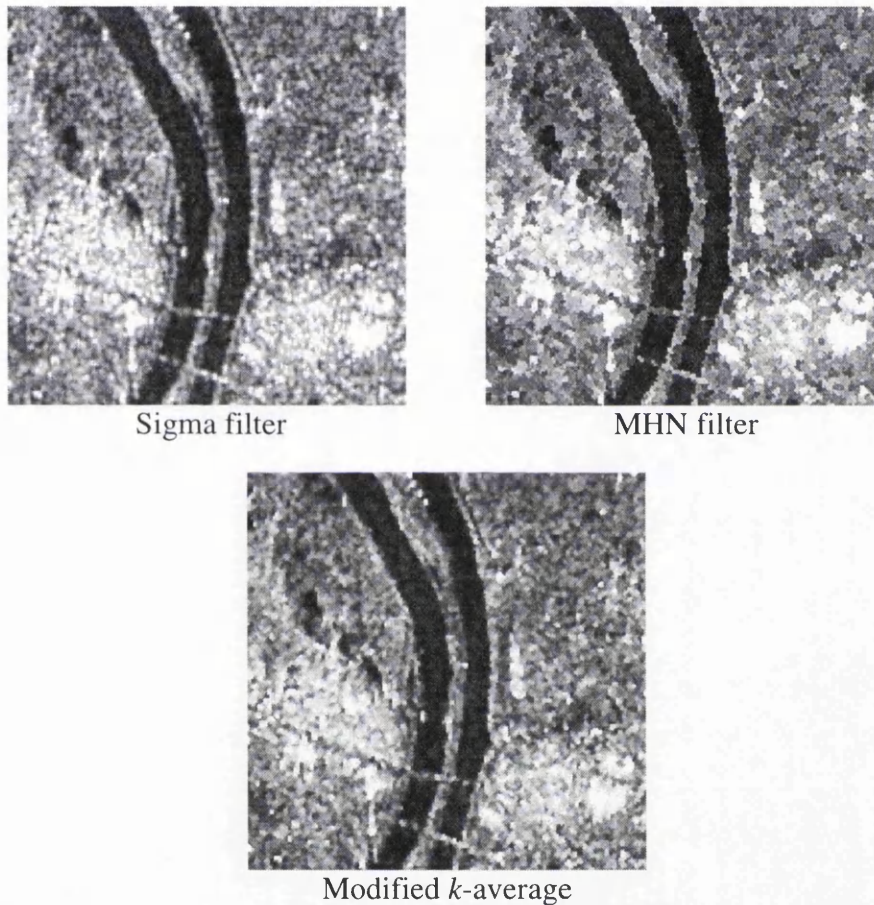


Figure 4.8 Sigma filter, MHN filter, and modified k -average filter applied to Beaucaire test image

Not surprisingly, these results are different to those produced by the previous two groups of filters. The sigma filter seems to have done a good job of removing speckle but retaining features, as has the modified k -average filter, although this one has removed a little more speckle. The most interesting results are those produced by the MHN filter, where the smoothed image has quite a granulated appearance. The speckle appears to have been all but eliminated, and the edges of features retained, but it is not easy to be sure whether this result is more useful than any of the others.

4.2.5 Improved traditional filters

This section describes improvements to traditional filters and the improved application of traditional filters, which have been developed in the course of this research.

One method of using standard smoothing filters to give better results is by applying them repeatedly to the same image. For a filter which does not remove very much speckle on its first application, a second, third or even fourth application may remove more speckle whilst still retaining features in the image. However, for some filters, after the first application the statistically independent nature of the pixels may be destroyed, in other words, the image becomes correlated. Therefore, for the subsequent applications of the filter, this has to be taken into account. The result is an iterative application of speckle reduction filters, which for some filters, such as the MAP filter, proves very successful. Further details can be found in McConnell and Oliver (1996).

A second method of improving speckle reduction filters is by applying more than one filter in combination. Although this will only work for a few types of filter, it does produce some very useful results. In particular, the combination of the median filter with two of the three adaptive, non-rigorous filters described above (sigma, MHN) as well as the k -average filter, was investigated. In each of these filters, some form of averaging is done on some of the pixels surrounding the pixel being processed. It was felt however, that the performance of each of these filters could be improved by replacing the averaging algorithm with the median algorithm, since the median algorithm gives better retention of features than the mean algorithm, but they are both able to reduce speckle by a similar amount (see table 4.1 in § 4.2.5). Therefore, three new filters are proposed here: the sigma-median filter, the MHN-median filter and the k -median filter.

In the traditional sigma filter, the output pixel value is given by the mean of all the pixels in a surrounding window which are within two standard deviations of the input pixel value. In the case of the sigma-median filter, the median of the pixels, rather than the mean is taken. Therefore, although the same number of pixels are used to generate the output value, pixels which differ greatly from the input pixel value have less influence on the result. Figure 4.9 shows the sigma filter and the sigma-median filter both applied to the Beaucaire test scene.

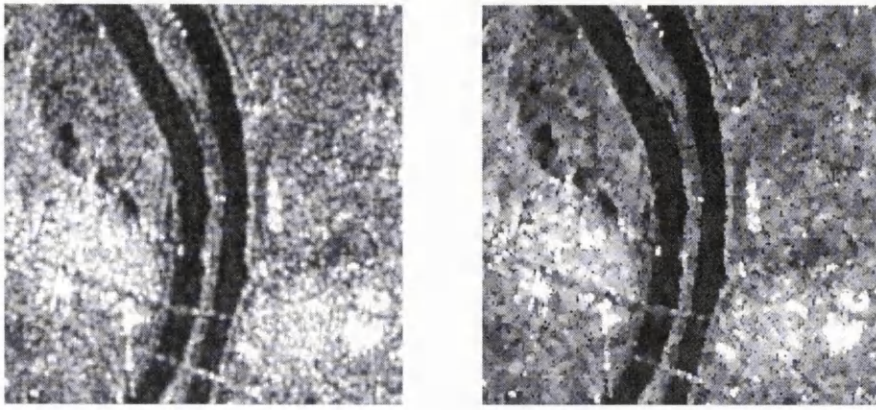


Figure 4.9 Sigma filter and sigma-median filter applied to Beaucaire test image

As with the sigma-median filter, the MHN-median filter generates its result using the median algorithm rather than the mean algorithm. The same argument applies as before: pixels whose values differ greatly from the input pixel value have a minimal effect on the output pixel value. Figure 4.10 shows the MHN filter and the MHN-median filter both applied to the Beaucaire test scene.

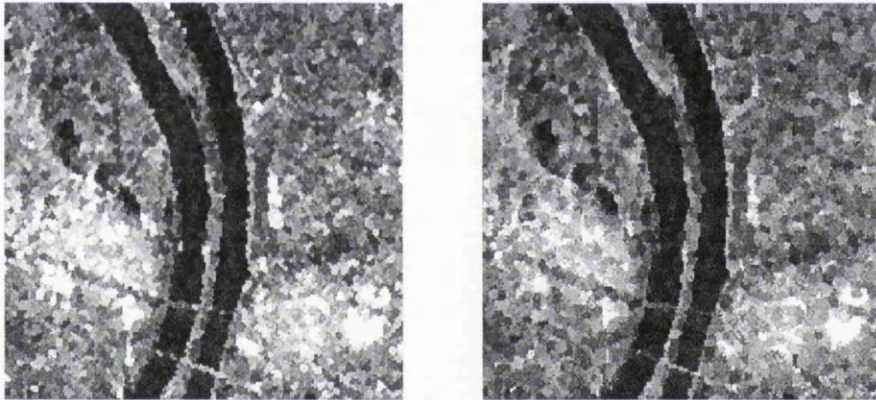


Figure 4.10 Traditional MHN filter and MHN-median filter applied to Beaucaire test image

The k -median filter works in the same way as the k -average filter, except that the median of the k nearest pixels is calculated rather than the mean. When applied using a 3×3 window, the median of the six pixels whose values are closest to the input pixel value is assigned to the output pixel. When a 5×5 window is used, the median of the seventeen pixels whose values are closest to the input pixel value is assigned to the output pixel. To keep in line with all the other filters described here, the k -median filter has been applied using a 5×5 window. Figure 4.11 shows the k -average filter and the k -median filter both applied to the Beaucaire test scene.

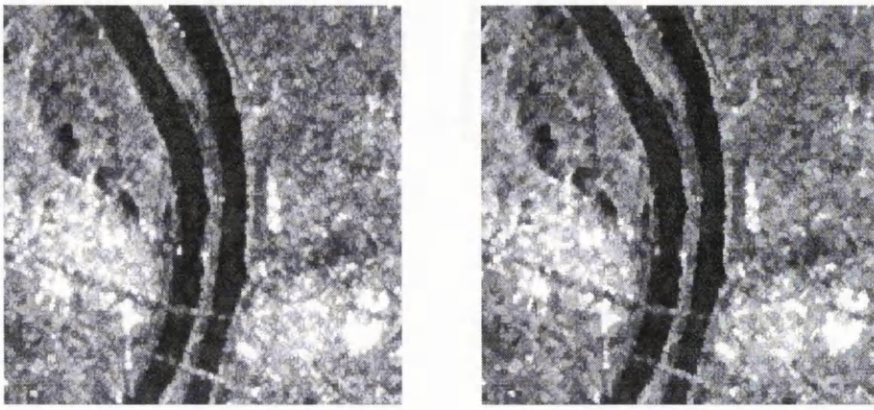


Figure 4.11 k -average filter and the k -median filter applied to Beaucaire test image

It is not easy to see whether or not each of these new filters gives an improved visual result compared to the corresponding original filter. In each case, just as much speckle appears to have been removed, but the features in the image have been affected differently, giving the different results. To evaluate the usefulness of these three new filters, they have to be compared with all the other filters discussed above. A comparison of all the filters described in this chapter is given in the next section, which introduces two parameters that can be used to compare the properties of the different filters.

4.2.6 Comparison of speckle reduction filters

One of the main difficulties with speckle reduction is not developing a good speckle reduction filter, but defining what is actually meant by the term ‘good’. The previous four sections have described a number of different speckle reduction filters. All of these filters operate quite differently, and give a wide range of results. The reason for trying to remove speckle from a SAR image is to increase the interpretability of the image and the detectability of features. Therefore, the two goals of speckle reduction filters, both of which must be achieved, are reduction of speckle and retention of features. Therefore when comparing different filters, these two attributes can be used as a measure of the quality of the filters.

A method of measuring the speckle reduction capability of filters, proposed by Crimmins (1986), is the speckle index. The speckle index is defined by the equation:

$$\text{Speckle index} = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N \frac{\sigma(m,n)}{\mu(m,n)} \quad (4.12)$$

where σ and μ are the variance and mean of pixels measured within a 3×3 sliding window passed across an image with dimensions of M and N . The speckle index gives a

measure of the amount of speckle in a particular SAR image. By measuring the speckle index both before and after a filter has been applied, it is possible to get some idea about how much speckle has been removed in the filtering process.

A measure of the retention of features during speckle filtering is much harder to quantify. McFarlane and Thomas (1984) proposed a simple, but not necessarily particularly robust method of assessing the retention of features using step edges in the image. They suggested locating a few step edges in the unfiltered image, and measuring their gradients. After the image has been filtered, the gradients are measured again and the ratio of the two gradients (before and after) for each edge will give an indication of the amount of smoothing of the edges that has taken place, and hence how well features have been retained. Narasimha Rao et al. (1995) used this principle to define a quantity for assessing the strength of edges: the edge strength index. A further variation is proposed here for determining how well edges have been retained after the application of a speckle reduction filter. In the Beaucaire test image, ten adjacent step edges located along the bank of the river were interactively selected (figure 4.12). The difference between the high DN of the riverbank, and the low DN of the river was measured for each of the ten step edges, and the results averaged. When repeated for the filtered image the ratio of the average step heights gives an indication of the amount of edge retention.

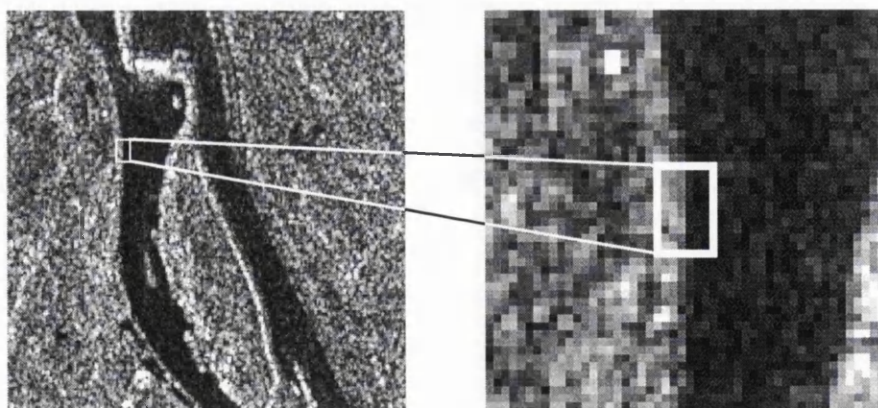


Figure 4.12 Location of step edges in Beaucaire image

Furthermore, it is useful to view the profile of this edge in three dimensions to understand visually how the smoothing algorithms affect the sharpness of the edge. Figure 4.13 shows the above step edge in the original image displayed in three dimensions: rows and columns are represented by the x - y plane, and the z -axis displays the DN values.

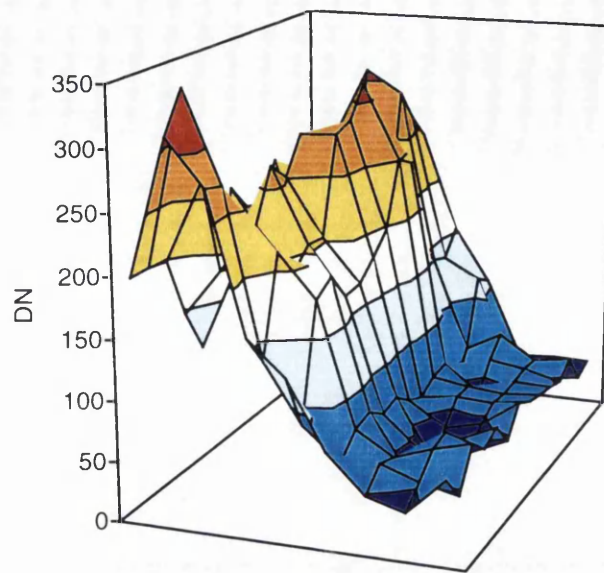


Figure 4.13 Perspective view of step edge in Beaucaire test image

These two measures were applied to a selection of smoothed SAR images generated from the speckle reduction filters described above. The results are shown in table 4.1. The relative speckle index was calculated by dividing the speckle index for the smoothed image by the speckle index of the original image, meaning that a value close to 1 represents low smoothing, whilst a value close to zero represents high smoothing. Similarly, for the edge retention indicator, a value close to 1 represents good retention of features, whilst a value close to zero represents poor retention of features. A combined parameter can be generated where the edge retention parameter for each filter is divided by the speckle index for that filter. The result is a parameter which gives an indication of the overall quality of the filter: the higher the value, the better the filter. All of the filters were applied using a 5×5 sliding window. They have been grouped together in table 4.1 according to type of filter.

Speckle reduction filter	Relative speckle index	Edge retention indicator	Edge retention/speckle index
Mean	0.1033	0.294	2.85
Plus mean	0.2114	0.569	2.69
Median	0.0969	0.777	8.02
Gaussian	0.2674	0.587	2.20
<i>k</i> -average	0.2861	0.922	3.22
Frost	0.4667	0.963	2.06
Lee	0.4637	0.981	2.12
Kuan	0.4569	0.838	1.83
MAP	0.1246	0.348	2.79
Sigma	0.2245	0.806	3.59
MHN	0.4129	1.169	2.83
Sigma-median	0.3724	1.126	3.02
MHN-median	0.3152	1.003	3.18
<i>k</i> -median	0.2488	1.065	4.28

Table 4.1 Speckle indices and edge retention indicators for various speckle reduction filters

With regards to which filters are best at removing speckle, the speckle indices in table 4.1 confirm what has already been seen in the above examples, namely that the simple filters remove the most speckle, closely followed by the MAP filter. The worst filters, in terms of the amount of speckle removed, are the rigorous adaptive filters (Lee, Kuan and Frost). The rest of the filters fall somewhere between these two extremes. This result is not really that surprising: simple filters smooth the image quite brutally, irrespective of whether any details need to be preserved, whereas rigorous adaptive filters pay great attention to removing only speckle and thus retaining a large amount of fine detail. It should be remembered that the speckle index is really just a measure of fine detail within the image: it makes no distinction between speckle and features. Therefore it is not possible to judge whether speckle or features, or both, are being removed from the image. It is for this reason that the edge retention indicator was introduced.

The edge retention indicator gives a good indication of the extent to which edges are smoothed during the filtering process. The worst culprits for smoothing edges are the mean and MAP filters, whilst four of the five non-rigorous adaptive filters actually enhanced the edge feature. The rigorous adaptive filters are also very good at retaining edges.

In order to decide which filters could be considered the best for removing speckle but retaining edges, the ratio of edge retention parameter to speckle index provides a useful guide. The median filter seems to be the best, and the Kuan filter the worst. All three of

the new proposed filters (sigma-median, MHN-median and k-median) give very good results, as does the sigma filter. These results should not be interpreted too rigidly since they can vary according to which test image is used, or how various parameters have been set, but they are confirmed to some extent by previous studies: the median filter is generally accepted as being one of the best filters available (Toll, 1985; Mueller and Hoffer, 1989), and the MHN filter has often been used as a prelude to feature extraction from Earth observation images (Newton et al., 1994; Morgado and Dowman, 1997; Vohra and Dowman, 1996).

Figures 4.14, 4.15 and 4.16 show the perspective view of the step edge after the application of selected filters.

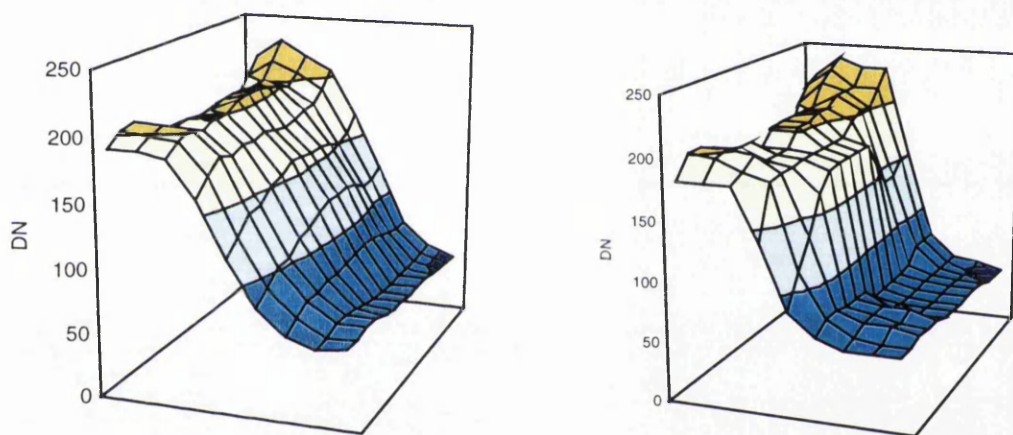


Figure 4.14 Perspective view of step edge after application of mean filter (left) and median filter (right)

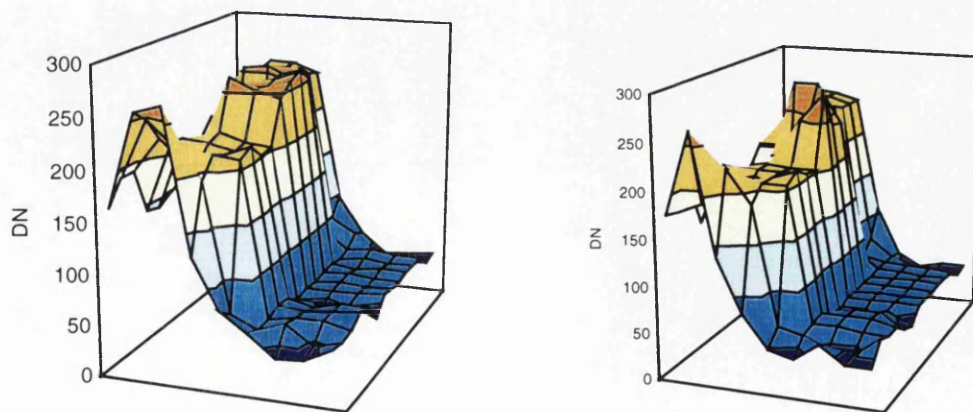


Figure 4.15 Perspective view of step edge after application of MHN filter (left) and MHN-median filter (right)

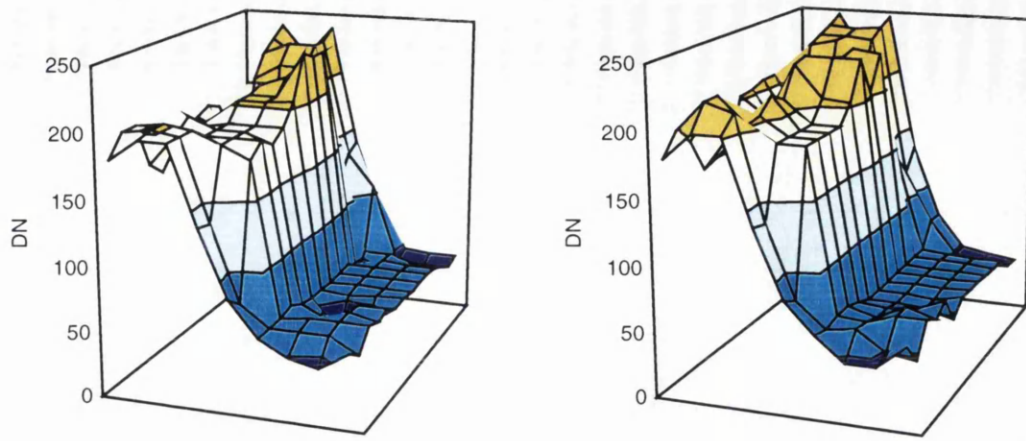


Figure 4.16 Perspective view of step edge after application of sigma-median filter (left) and k -median filter (right)

It can clearly be seen that the mean filter smooths the edge more than any of the filters. The steep slope is completely wiped out, leaving quite a gentle gradient from the bank to river. In the image this would mean the complete loss of distinct edges. The median filter smooths the bank and the river very well, but retains a lot of the steepness of the edge. The other four filters show varying degrees of smoothing and edge retention. The MHN and the sigma filters smooth the features slightly more than the corresponding MHN-median and sigma-median filters, as would be expected, but all seem to retain the steepness of the slope as well as each other.

This concludes the comparison of standard speckle reduction filters. Further comparisons can be found in McFarlane and Thomas (1984), Mastin (1985) and Dare and Dowman (1996).

4.2.7 Conclusions

From the above discussion of speckle reduction, it can clearly be seen that development of smoothing filters which do not destroy image information is not a trivial problem. So much work has been carried out over the years on this subject that it is not possible to describe it all here. Hence, the above discussion has limited itself to the main speckle reduction techniques which are generally accepted as being useful in SAR image interpretation.

Four groups of speckle reduction filters were introduced: simple filters, rigorous adaptive filters, non-rigorous adaptive filters and modified traditional filters. Each group gave quite different types of results: simple filters removed speckle well, rigorous adaptive filters preserved features well and the other two groups fell somewhere in

between. The aim of this study is to select those filters which best aid feature extraction for image registration. However, before feature extraction has been investigated it is difficult to say at this stage exactly which filters should be selected, although it is possible to describe general properties of these filters, based on the types of features will need to be extracted: namely polygonal features that are recognisable in both of the images. These features must therefore have closed boundaries and they must not be too small, since small features are difficult to match. Therefore the ideal speckle reduction filters for this application can afford to lose fine detail, so long as features with strong edges are retained. This eliminates the simple filters (too much smoothing, therefore weak edges) and rigorous adaptive filters (not enough smoothing, therefore too many small features), hence leaving the original and modified non-rigorous adaptive filters. The best performers were the sigma filter, the MHN filter, the sigma-median filter, the MHN-median filter and the k-median filter. These filters will be used in the next stage of the processing, along with the median filter, which gave consistently good results.

4.3 Preprocessing of SPOT data

4.3.1 Introduction

Unlike SAR data, there is very little radiometric preprocessing that actually needs to be performed on SPOT data before it is used in the feature extraction process. Since the signal to noise ratio is usually good, noise does not obscure features and no smoothing is required. However, to improve the success of SPOT feature extraction algorithms, application of particular smoothing filters may be necessary.

Before any algorithms were applied to the SPOT images used in this study, the contrast was improved by adjusting the image histogram. A histogram normalization was performed, making features become much more distinct from their backgrounds. It was anticipated that this would improve the effectiveness of feature extraction algorithms without destroying any image information.

4.3.2 Smoothing of SPOT images

The smoothing filters that may aid the extraction of features are those which enhance the strengths of edges in the image, but do not destroy too much image information. From the study of SAR speckle reduction, it was decided to test three smoothing algorithms: the median filter, the MHN filter, and the MHN-median filter. The results of applying these algorithms to SPOT images are shown in figure 4.17 below.



No smoothing



Median filter



MHN filter



MHN median filter

Figure 4.17 Smoothing algorithms applied to SPOT image of Beaucaire

The median filter (applied here using a 3×3 kernel) apparently does not change the original image significantly. The structure of the features has been preserved, and a small amount of noise has been removed. The MHN filter has much more of an influence than the median filter. The result is that features appear sharper, but it is clear that information has been lost, especially in the urban areas just below the centre of the image. This loss of information is not considered to be a problem since only very small features have been affected badly – the larger features are still very clear. The MHN-median filter produces very similar results to the MHN filter – so much so in fact that it is very difficult to detect any visual differences between the two resulting images. As a result it was decided to use only the median and MHN filters in conjunction with the feature extraction algorithms.

As a final point it is interesting to note that the MHN filter and the MHN-median filter produce acceptable results for both images that are very noisy (SAR images) and those which are not so noisy (SPOT images). It is only when a lot of noise is present that the differences between these two filters is noticeable. This confirms that although the MHN-median filter removes more noise pixels than the MHN filter, it still preserves a significant amount of the information in the image.

4.3.3 Conclusions

The higher radiometric quality of SPOT panchromatic images compared to ERS SAR images means that relatively little preprocessing is required before features are extracted. Two algorithms were selected to be used with the feature extraction algorithms discussed later in this thesis. In addition, the contrast of the SPOT images was improved using a histogram normalization.

4.4 Discussion

This chapter has described the preprocessing of the SAR and SPOT images used in this study. Due to the radiometric quality of the SAR images comparatively more processing was needed than for the SPOT images.

With an understanding of SAR imaging, it has been shown that it is possible to construct a mathematical model of speckle, and hence develop rigorous speckle reduction algorithms. These algorithms have been shown to be very efficient at removing speckle from SAR images whilst retaining many of the features within the image. A number of non-rigorous filters were also introduced, which in some cases performed just as well, even though they did not use a rigorous model of speckle. Thus, for speckle reduction, although an understanding of SAR imaging is useful, it is not necessarily essential. Interestingly, it was found that the speckle reduction algorithms which did not rely on a rigorous mathematical model gave better results for the purposes of this study.

An important part of this chapter is the proposal of three new speckle reduction filters: the sigma-median, the MHN-median and the k -median. They do not attempt to estimate the radar backscatter like some of the rigorous filters, but instead they manipulate the image to give a better visual result, which for the purposes of this study is perfectly adequate. Although they are not necessarily fundamentally ground breaking, they are new and do produce good results compared to the other filters tested.

Compared to the preprocessing of the SAR data, the SPOT preprocessing is very simple. The only algorithms which can be usefully applied to the SPOT data are those which enhance features while only slightly smoothing the image. The filters that fall into this category are the median filter, the MHN filter and the MHN-median filter. However, only the median and MHN filters will be used in later chapters – the results produced by the MHN-median filter were too similar to the MHN results to be worth considering.

This concludes the overview of preprocessing techniques used in this study. The algorithms introduced here are now used in the next few chapters in conjunction with feature extraction algorithms.

Chapter 5 FEATURE EXTRACTION

5.1 Introduction

Chapter 2 introduced the subject of image registration based on feature matching, and proposed a system requiring the extraction and matching of both patches and edges. This chapter now describes in detail the principles behind feature extraction for automatic image registration, and the practicalities involved in applying feature extraction algorithms to different images. The subject of feature extraction, along with feature matching (described in the next two chapters) comprise the most difficult problems that have to be solved for automatic feature based multisensor image registration.

The ability to successfully match features from different images depends very much on the type, quantity and quality of the extracted features. If the features extracted are of the wrong type, poor quality, or too few, then it will be extremely difficult, if not impossible, to match corresponding features and perform the registration. Therefore a lot of time and effort has been put into the development of feature extraction algorithms to help ensure there are plenty of suitable features extracted from each image to be matched.

The three themes central to feature extraction for image registration are:

- ❑ similar features have to be extracted from each image, even if the images have very different spectral properties; but,
- ❑ not every single feature needs to be extracted from both of the images; and,
- ❑ not all extracted features need to be matched.

It is essential that similar features are extracted from the pair of images being registered, and that those features are well distributed across the images. Once the features have been extracted, they have to be matched. If field boundaries, for example, are extracted from one of the images, then field boundaries must also be extracted from the other

image. This is not a problem if the images were acquired by the same type of sensor, or by different sensors operating in the same part of the electromagnetic spectrum, but if the images are very different in nature (as is the case with SAR and optical data) then it proves more difficult to extract similar features. Therefore, to improve the chances of extracting similar features from different images, a range of feature extraction algorithms have to be used.

Unlike other applications of feature extraction, such as landcover classification, it is not necessary in feature based image registration to extract every single feature in the image. For a pair of relatively small images (say, for example, 512×512 pixels) a minimum of four well distributed matched patches would be sufficient to perform the first stage of the image registration procedure. To do this would mean extracting exactly the same four features from each image, which although not impossible, is quite unlikely. Therefore as many patches as possible should be extracted to ensure a large redundancy of features.

In the same way as it is not necessary to extract every feature from the image, it is also not necessary to match every feature that has been extracted. Even if hundreds of features have been extracted, four well distributed matched features are sufficient to perform the registration. However, if more features can be matched, a better result would be achieved since again there would be greater redundancy.

To summarize feature extraction in the context of image registration, it is reasonable to say that redundancy is a key factor. Although not all features need to be extracted from the image, as many as possible should be extracted to ensure a redundancy of features, and although not all the extracted features need to be matched, as many as possible should be matched to ensure a redundancy of matches. It is this redundancy that will eliminate blunders from the automatic image registration procedure, and hopefully lead to a more accurate result.

A final comment should be made about the types of features used in the automatic registration procedure. Although linear and areal feature extraction techniques are discussed in this chapter, there are other types of features which could be used for feature matching, but which have not been included in this study. Two principal ones are point features and wavelets (Djamdji et al., 1993). The reason for not using point features was given in chapter 2: point features extracted from multisensor imagery cannot be matched. It has been shown by many authors that wavelets can be

successfully used for single sensor image registration (Fonseca et al., 1997; Devore et al., 1997). However, the possibilities of using wavelets for multisensor image registration are less well researched. A quick study of wavelets seemed to show that they would not necessarily be any better for multisensor image registration than linear or areal features, so they are not considered in this study. However, further research is required in this area.

5.2 Edge extraction

5.2.1 Introduction

This section describes the various procedures investigated in the course of this study to extract edges from optical and microwave remotely sensed imagery. The two main types of edges in images are boundaries between two different regions on the ground (such as boundaries between different agricultural fields), and long, thin, continuous features (such as roads or railways). In this study both of these types of features have been grouped together under the title of edges, even though strictly speaking the second group can be thought of as linear features rather than edges. The following sections describe a range of methods for detecting edges using derivative operators, an algorithm for detecting edges in SAR images (SCANEDGE) and other techniques for extracting linear features specifically.

Extraction of edges is much simpler than extraction of linear features since an edge can be modelled much more easily. An edge in an image is just a boundary between two regions of different grey level and it does not necessarily represent a real feature on the ground. However, the strength of an edge can easily be determined, and it is reasonable to assume that if the edge is strong it does actually represent a real feature. Therefore, by simply thresholding an edge image it should be possible to eliminate false edges. In contrast, linear features (such as those which represent roads or railways) are far more complicated. Their grey level must be different from the background for them to be visible, but this grey level may vary as the background varies. The development of an algorithm which can recognise the complexities of a linear feature and extract it from the background is not at all trivial, and beyond the scope of this work. Therefore, in this study, the extraction of edges is pursued, but the extraction of linear features is not. However, a description of linear feature extraction techniques has been included.

5.2.2 Gradient operators for edge detection

The most efficient way of detecting edges in images is by using derivative spatial filters (Gonzalez and Woods, 1992). If the image is represented by the function $f(x, y)$ then the first derivative of this function is represented by the vector ∇f , where

$$\nabla f = \mathbf{i} \frac{\partial f}{\partial x} + \mathbf{j} \frac{\partial f}{\partial y} = \mathbf{i} \nabla f_x + \mathbf{j} \nabla f_y \quad (5.1)$$

The magnitude and direction of this vector are given by

$$\text{Magnitude, } |\nabla f| = \sqrt{(\nabla f_x)^2 + (\nabla f_y)^2} \quad (5.2)$$

$$\text{Direction, } \theta(x, y) = \tan^{-1} \left(\frac{\nabla f_y}{\nabla f_x} \right) \quad (5.3)$$

In practical terms, the images are differentiated by applying a kernel to the image with elements that approximate the above equations. Consider the 3×3 kernel:

a_1	a_2	a_3
a_4	a_5	a_6
a_7	a_8	a_9

(5.4)

An approximation to the first derivative at a_5 in the x (horizontal) and y (vertical) direction is given by the partial derivatives:

$$\frac{\partial f}{\partial x} \approx (a_1 + a_4 + a_7) - (a_3 + a_6 + a_9) \quad (5.5)$$

$$\frac{\partial f}{\partial y} \approx (a_1 + a_2 + a_3) - (a_7 + a_8 + a_9) \quad (5.6)$$

These expressions can be represented by two kernels (shown below in equation 5.7) which, when convolved with the image, give results that can be combined to determine the magnitude and direction of edges at every pixel location.

1	0	-1
1	0	-1
1	0	-1

1	1	1
0	0	0
-1	-1	-1

(5.7)

These particular kernels are known as the Prewitt operators. There are many other types of operators which approximate the partial derivatives in equations 5.1, and therefore enhance the edge information in the images. The Sobel and Frei-Chen operators are shown below.

Sobel:

1	0	-1
2	0	-2
1	0	-1

1	2	1
0	0	0
-1	-2	-1

(5.8)

Frei-Chen:

1	0	-1
$\sqrt{2}$	0	$-\sqrt{2}$
1	0	-1

1	$\sqrt{2}$	1
0	0	0
-1	$-\sqrt{2}$	-1

(5.9)

As well as enhancing edge information in images, operators such as those described above also enhance noise which can lead to the detection of false edges. Unlike the Prewitt operator, the Sobel and Frei-Chen operators introduce a smoothing effect as well as an edge enhancement effect, which gives a slightly less cluttered image, but the edges are very slightly less distinct. For this study it was decided that the Sobel operator would be used since it gives the most smoothing. Figure 5.1 shows a test image and the corresponding edge strength image generated using the Sobel operator.



Figure 5.1 Original and edge strength images generated using Sobel operator
(© Playboy Inc.)

The edge strength image in figure 5.1 above shows how the Sobel operator highlights strong edges well (such as those around the hat) but weak edges are more difficult to

make out (the line of the nose and mouth). Before these edges can be used in the matching procedure described in chapter 7, a lot of post-processing has to take place; this is described below in § 5.2.5.

Edge enhancement is not limited to first order derivatives: second order derivatives can also be used. The two dimensional Laplacian of the image function $f(x, y)$ is given by

$$\nabla^2 f(x, y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (5.10)$$

The most common way of determining the second derivative of a digital image function is by convolving the image with the operator:

0	-1	0
-1	4	-1
0	-1	0

(5.11)

However, problems can occur when applying this operator for two reasons: it will produce double edges, and it tends to enhance noise pixels as well as edge pixels. The production of double edges can be explained by examining the profile of an edge (figure 5.2).

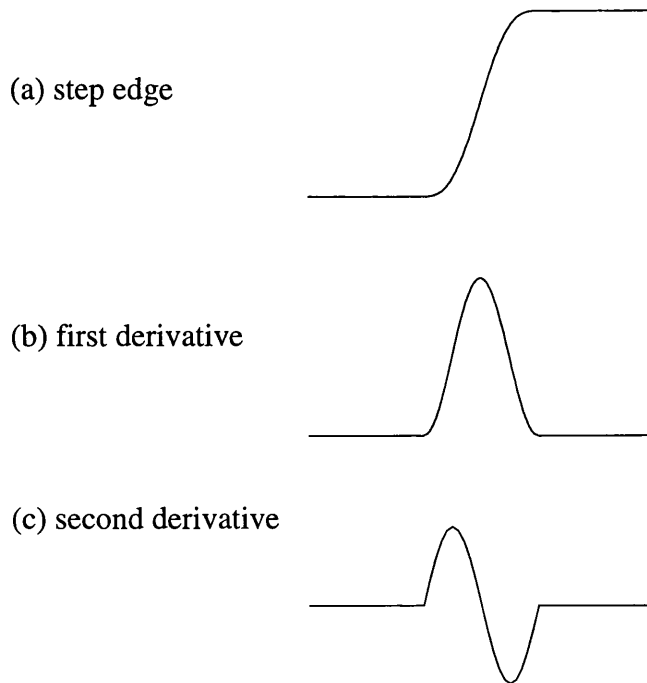


Figure 5.2 Step edge with first and second order derivatives

Figure 5.2 (a) shows a typical step edge, figure 5.2 (b) shows the first derivative of this step edge, and figure 5.2 (c) the second derivative. The first derivative has one peak but the second derivative has a peak and a trough. The magnitude of the second derivative therefore has two peaks. However, the second derivative can be used to locate the edge by examining the location of the zero-crossing. The procedure proposed by Marr and Hildreth (1980) suggests firstly smoothing the image with a Gaussian filter to limit the rate at which intensities change, and then identifying intensity changes (edges) by finding the zero-crossing of the second derivative. Hence the image is convolved with the Laplacian of Gaussian. The Gaussian takes the form:

$$G(r) = \exp\left(\frac{-r^2}{2\sigma^2}\right) \quad (5.12)$$

where $r^2 = x^2 + y^2$ and σ is the standard deviation.

$$\text{Laplacian of Gaussian} = \nabla^2 G = \left(\frac{r^2 - \sigma^2}{\sigma^4}\right) \exp\left(\frac{-r^2}{2\sigma^2}\right) \quad (5.13)$$

In practical terms, the Laplacian of Gaussian operator is applied by setting a width parameter w , where

$$w = 2\sqrt{2}\sigma \quad (5.14)$$

Figure 5.3 shows the result of applying the Laplacian of Gaussian operator to an image with two different values for w .

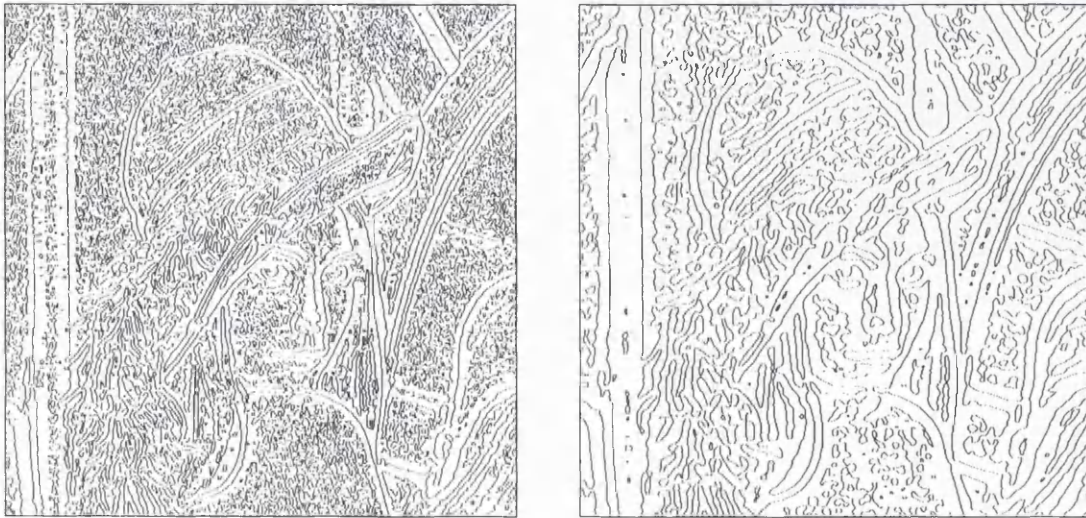


Figure 5.3 Zero crossings in test image, $w=5$ (left) and $w=9$ (right)

The edge image on the left of figure 5.3 ($w=5$) has clearly detected more edges than the image on the right ($w=9$). This is because with a larger value for the parameter w , there is much more smoothing in the image, and more of the weaker edges are smoothed out. However, the drawback of this is that the edges which have been extracted are much more generalized and do not necessarily reflect the actual edges in the original image quite as accurately as the edges extracted by using a lower value of w . Although both of the above images could be said to have some edges which truly represent real features in the original image, there are also a lot of edges present which represent only very weak features. It will be necessary to remove these types of edges since they will complicate and slow down the matching process, and possibly lead to false matches. Unfortunately, the principal method of removing weak edges from an image, non-maximal suppression (described below), cannot be used on edge images generated using second derivatives since there is no edge strength or edge direction associated with these images. Therefore, simple clutter removal algorithms (described at the end of this chapter) have to be used, even though they will not necessarily produce good results. The post-processing of edge images in order to improve feature matching results is a subject which is discussed in § 5.2.5.

5.2.3 The SCANEDGE algorithm

A method which was used for extracting edges from SAR images was the SCANEDGE algorithm supplied as a part of the commercial software package Caesar, produced by N. A. Software Ltd., UK (Caesar, 1996). The algorithm detects edges using a combination of a scanning window and a scanning edge (Oliver et al. 1995). The user is able to set a parameter, p , which controls the probability of detecting false edges. Figure 5.4 shows edges extracted from the Entressen SAR image using three different values of p (0.01, 0.001 and 0.0001).

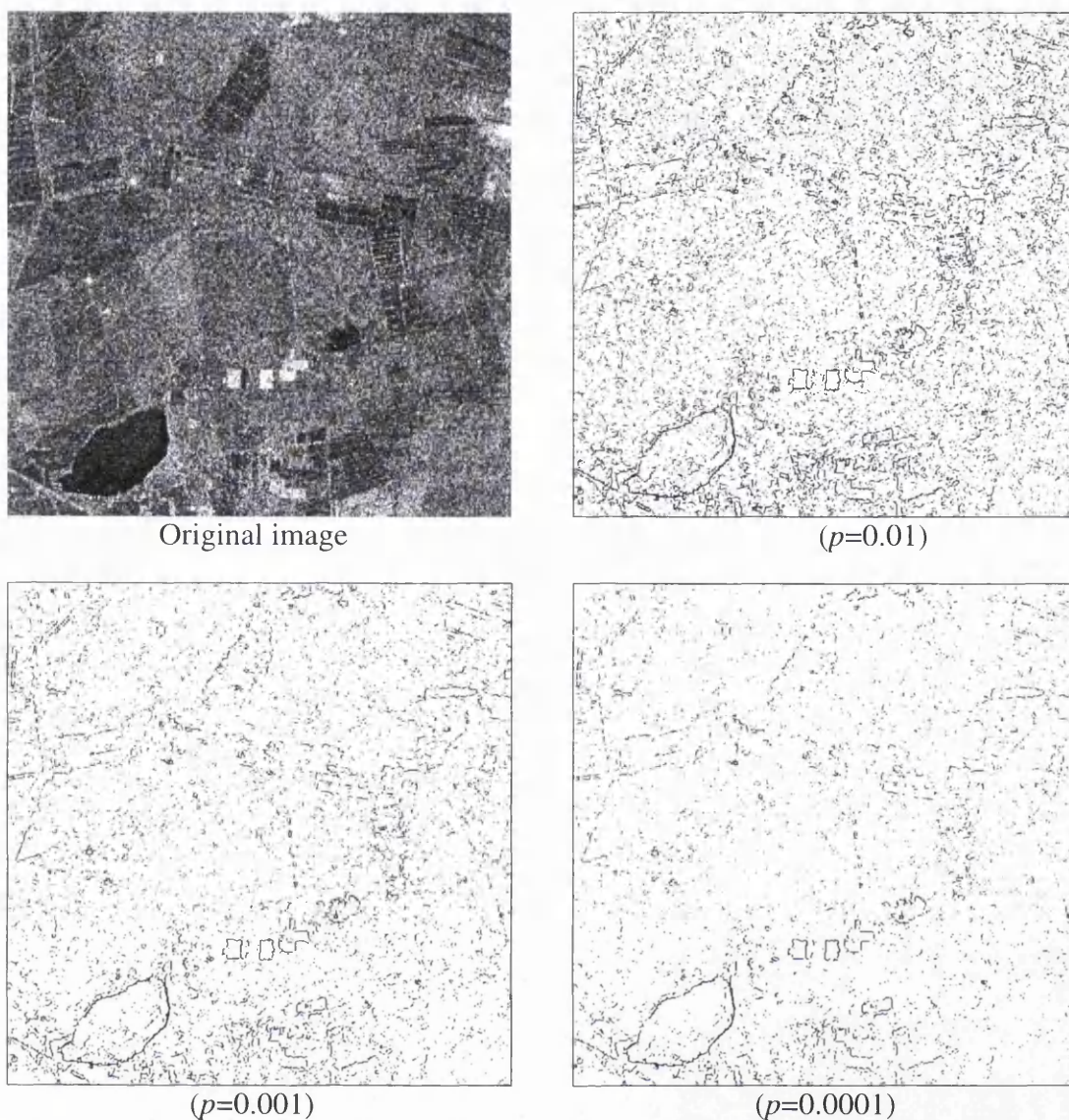


Figure 5.4 Edges extracted from Entressen images using SCANEDGE with three different values of p

The SAR image used to illustrate the SCANEDGE algorithm would be a challenge for any edge detector, even though it is a quite typical of SAR data. Although the human observer is able to clearly see field boundaries, there are really only a few features which stand out enough from the background to be detected by this algorithm. Those features are the lake in the bottom left corner of the image and the group of very bright rectangular objects near the centre. These features have been extracted well in each of the four different examples above, but where lower values of the parameter p have been used, far more edges have been extracted. However, it is not clear whether these edges represent real features on the ground, or whether they are false alarms. It would be expected that the more continuous the edge feature is, the more likely it is to represent a feature on the ground. Therefore, a careful balance between choice of the parameter p and clutter removal algorithms should lead to the best result for feature matching.

5.2.4 Other methods of linear feature extraction

In the course of this study, other linear feature extraction algorithms were tested, which although ^{they} did not produce good results at the time, have been included in this thesis due to their relevance to the subject. Both of the techniques described below would need more research before they could be used successfully in the feature based image registration procedure, but even so they are potentially very useful algorithms.

One of the linear feature extraction algorithms tested but not pursued was the Hough transform (Hough, 1962). A selection of points in an image $f(x, y)$ will all lie on a straight line if they satisfy the equation:

$$y = ax + b \quad (5.15)$$

where a and b are parameters representing the gradient and y-intercept of that line. Rewriting this equation as:

$$b = -ax + y \quad (5.16)$$

and substituting the x and y values of a point (x_k, y_k) in image space will yield a straight line in the ab plane (parameter space) with the equation:

$$b = -ax_k + y_k \quad (5.17)$$

If two points are collinear in image space, they will translate into two intersecting lines in parameter space. The point of intersection in parameter space represents the gradient and intercept of the line joining the two points in image space. Any further points which lie on the same line in image space will also translate into straight lines in parameter space which will all intersect at the same location. Therefore, by examining intersecting points in parameter space, it is simple to determine the parameters of the line passing through collinear points in image space. However, if those points lie along a line which is almost vertical, a and b approach infinity. Therefore, an alternative method of describing a straight line which avoids infinities can be used:

$$x \cos \theta + y \sin \theta = \rho \quad (5.18)$$

where ρ is the length of the normal to the line, measured between the line and the origin, and θ is the angle between the normal and the x -axis. The procedure for determining gradient and intercept is exactly the same, except that now lines are plotted in the $\rho\theta$ plane rather than the ab plane.

In practical terms, the Hough transform can be difficult to implement for two reasons: detected lines have infinite length, and a linear feature in an image is not represented by a series of points. To limit the length of detected lines, the image must be split into tiles, and each tile processed individually. To convert the linear features in the image to a series of points, the image must be preprocessed. It is this preprocessing step that needs further research before the Hough transform can be used to extract linear features from remotely sensed images in a robust manner. An experiment was performed with a SAR image of a region of Southern France near the town of Lançon. This image was thresholded and filtered for clutter before linear features were extracted with the Hough transform. Four processing steps are shown in figure 5.5.

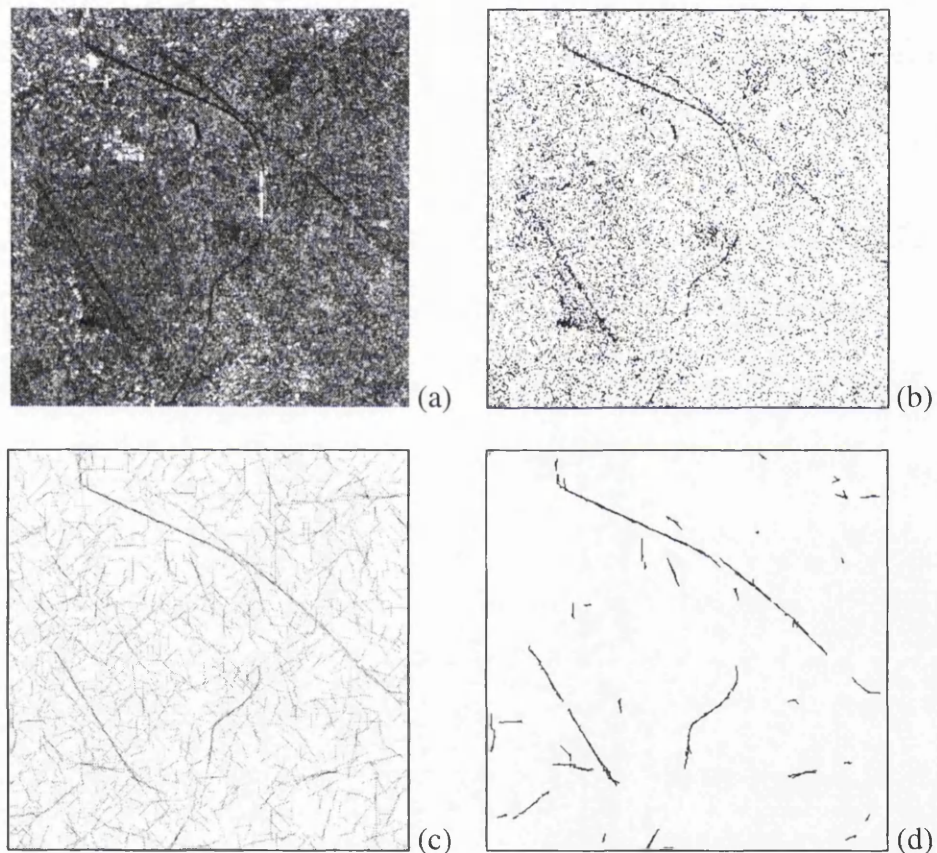


Figure 5.5 Four processing steps in Hough transform procedure

Figure 5.5 (a) shows the original Lançon SAR image, and figure 5.5 (b) shows the result of preprocessing, where the image has been thresholded and filtered for clutter. Figure 5.5 (c) shows the result of applying the Hough transform to a tiled version of the image, and in figure 5.5 (d) this image has been thresholded to separate the strong linear features from the weak ones. The resulting linear features can be seen to coincide well with those in the original image, but a lot of features in the image have been lost in the extraction process. It seems that the Hough transform is a very powerful tool for linear

feature extraction, even from an image where there is a lot of noise, but further research is required if this is to become a reliable method of feature extraction.

A second linear feature extraction technique which was investigated but not followed up is the use of line enhancement masks, which when convolved with the image, enhance linear features (Gonzalez and Woods, 1992). The operators used to enhance lines are:

$$\begin{array}{cc}
 \begin{array}{|c|c|c|} \hline -1 & -1 & -1 \\ \hline 2 & 2 & 2 \\ \hline -1 & -1 & -1 \\ \hline \end{array} & \begin{array}{|c|c|c|} \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & 2 & -1 \\ \hline \end{array} \\
 & (5.19) \\
 \begin{array}{|c|c|c|} \hline 2 & -1 & -1 \\ \hline -1 & 2 & -1 \\ \hline -1 & -1 & 2 \\ \hline \end{array} & \begin{array}{|c|c|c|} \hline -1 & -1 & 2 \\ \hline -1 & 2 & -1 \\ \hline 2 & -1 & -1 \\ \hline \end{array}
 \end{array}$$

If the first operator is convolved with an image where a horizontal line is present, the operator will enhance this line relative to the background. The result of this operator will be greater than the results from the other operators, since they will tend to highlight lines in other directions. Therefore, by applying all of these operators to an image, and comparing the results from each, it is possible to enhance linear features. Tests showed that although this produced interesting results, it did not provide the output required for feature matching. However, use of these operators with an algorithm such as the Hough transform could possibly lead to much better linear feature extraction, but further investigation is required.

5.2.5 Post-processing of edge images

After edges or linear features have been extracted, they must be processed in order to highlight the important features and suppress the unimportant ones. Figure 5.6 shows the result of applying the Sobel edge detector to the SPOT image of Istres.

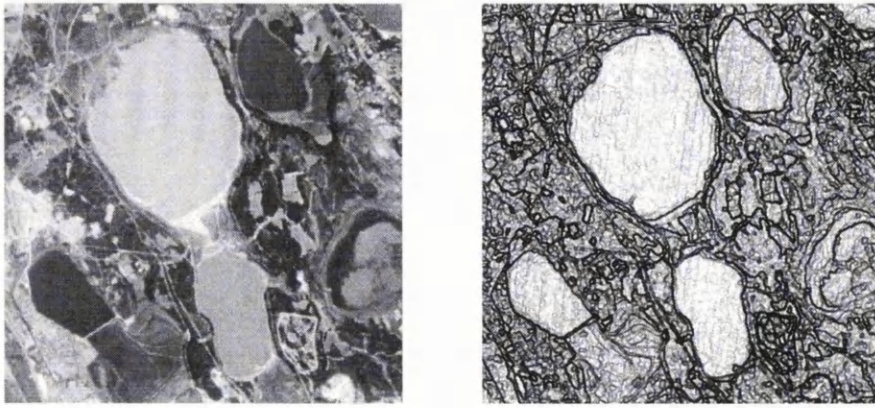


Figure 5.6 SPOT image of Istres (left) and Sobel edge strength image(right)

Although there are quite a few very strong edges in the image, there are also a lot of weak edges. In order that this result can be used efficiently in the matching procedure, the weak edges need to be removed. The technique used to do this is non-maximal suppression (Canny, 1986), but the particular method described here was proposed by Lewis (1988). The distinction of whether an edge pixel strong or weak is determined by examining neighbouring pixels. An edge pixel is defined as strong if its strength is greater than the two pixels which lie on a line parallel to the gradient of the edge in question and on either side of it (see figure 5.7).

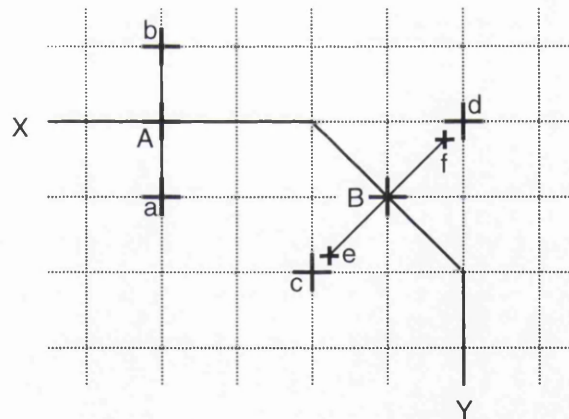


Figure 5.7 Non-maximal suppression proposed by Lewis (1988)

In figure 5.7, the edge follows a path from X to Y . Consider the pixel at point A . To decide if this pixel is strong it is compared with the pixels values at points a and b . If the strength of A is greater than both a and b , then it is considered strong. Now consider the pixel at point B . Rather than comparing the strength of B with the strengths of points c and d (as suggested by Canny, 1986), it is actually compared with the strengths at points e and f . These two points are located such that the length of ef is the same as ab , and e and f lie on a line parallel to the gradient of the edge at point B . The strengths at e and f are determined from a linear interpolation of the four surrounding pixel values. The result of removing weak edge pixels in this way is shown in figure 5.8.

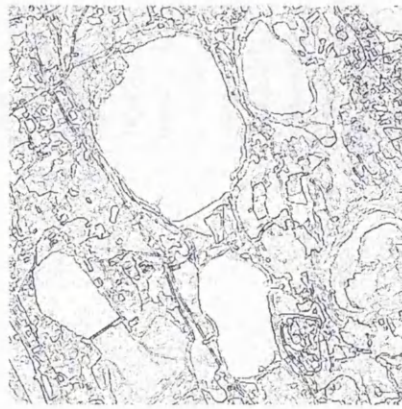


Figure 5.8 Weak edges suppressed using non-maximal suppression

There are still a large number of edges in the image so further processing is required: the edge suppressed image must be thresholded at some user defined value. The result is shown in figure 5.9.

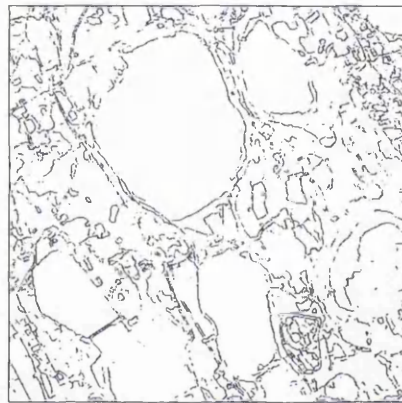


Figure 5.9 Thresholded edge image

It is essential that the edge image is thresholded so that weak edges can be removed, but selection of the threshold is very difficult since there is no obvious way of deciding where the boundary between strong edges and weak edges should lie. A method of trial and error has been employed here, but further research is necessary to formulate a more robust method.

5.2.6 Conclusions

This section has introduced some different methods of extracting edges. The list is by no means conclusive; it merely represents some algorithms which are useful for extracting edges for automatic image registration. Although the term ‘edge extraction’ has been used to describe all the algorithms here, they could really be separated into two categories: edge detection and linear feature extraction. Edge detection relies on differentiating the image either once (first derivative edge detection) or twice (second derivative edge detection). First derivative edge detection produces both an edge strength and edge direction image, whereas second derivative edge detection produces

an edge location image. Post-processing of both these images is necessary to prepare the images for registration by removing clutter.

The linear feature extraction algorithms described in this section include line detection kernels, the Hough transform and a commercially available edge detection routine for SAR images (SCANEDGE). Of these algorithms only SCANEDGE was tested fully. The Hough transform was shown to be a very useful tool for extracting straight linear features, but further research is required to ensure that consistently reliable results can be achieved. Line detection kernels were tested but found to not produce useful results for the purposes of this study.

Of all the edge extraction techniques described here, the only ones which can be used in conjunction with smoothing and speckle reduction algorithms introduced in the previous chapter are the first derivative edge detectors. The second derivative edge detector, the LoG operator, cannot be used in conjunction with a smoothing algorithm since it already incorporates one. The SCANEDGE algorithm relies on the DNs of the SAR image being uncorrelated, but this is only the case for raw images that have been processed in the correct manner. Application of a speckle reduction algorithm will lead to the pixel values being correlated and hence the SCANEDGE algorithm may produce unreliable results. Examples of results obtained when the Sobel operator is used with smoothing and speckle reduction algorithms are shown in chapter 7 when edge matching is discussed.

In summary some fairly basic linear feature extraction techniques have been tested and they have been found to produce results good enough for the purpose of image registration. Implementation of these results will be discussed in chapter 7.

5.3 Patch extraction

5.3.1 Introduction

This section describes the three techniques used in this study (and one other not used) to extract areal features from satellite imagery. A patch is any feature in an image with a two dimensional spatial extent. Although the extent of a particular feature in an image is easy for the human visual system to determine, it is much more difficult to train a computer to recognise where features begin and end. Therefore, the algorithms described below are concerned with recognizing a feature and its spatial extent, and then separating that feature from the surrounding features and background. A patch can be recognized by its colour, texture or boundary (or a combination of these three). Of the

four techniques described below, thresholding recognizes features according to grey level (i.e. colour), homogeneous patch extraction according to texture (or lack of it), and segmentation according to both grey level and texture. The method of extracting patches by converting boundaries to regions was not tested in this thesis, but is an important tool for feature extraction which is closely related to the other methods described, and so has been included here for completeness. In the very active discipline of computer vision, many techniques for extracting areal features from images have been developed, and continue to be developed, but to describe them all here would digress to far from the principle topic of this thesis (image registration). So long as the areal feature extraction algorithms introduced here provide useful features for image registration, then a full description of every other type of feature extraction algorithm is not necessary.

Of the algorithms described below, only three benefit from being used in conjunction with smoothing and speckle reduction algorithms. Automatic thresholding makes use of smoothing in order to enhance the image histogram (described in § 5.3.2), and the optical segmentation algorithms, REGSEG and OPTISEG, make use of the MHN filter to enhance the edges of features in the SPOT images. The reasons why preprocessing can or cannot be used with particular feature extraction algorithms are given in each case.

5.3.2 Thresholding

Thresholding is the simplest method of extracting areal features from digital images, but it is not really recognized as a reliable feature extraction tool. However, because of its relative simplicity, and success in other automatic image registration systems (Morgado and Dowman, 1997; Vohra et al. 1996) it was felt that it should be included in this study. Since a digital image consists of pixels with discrete grey values, it is possible to select a particular grey level (the threshold) below which all the pixels are set to zero, and above which all the pixels set to 255 (for an 8 bit image). Figure 5.10 shows an 8 bit greyscale image (i.e. DN values from 0 to 255) which has been thresholded at grey level 112.



Figure 5.10 Original greyscale image (left) and image thresholded at grey level of 112 (right) (© Playboy Inc.)

The result of the thresholding is that the image has been split into black and white areas corresponding to light and dark regions in the original image. If a particular feature is very light in colour and has a dark background, thresholding provides a means to separating that feature from its background. However, there are two particular difficulties associated with feature extraction using thresholding:

1. If the grey level difference between features and background varies across the image, a global threshold (i.e. thresholding applied to the whole image) will not extract all the required features.
2. Even in local regions (i.e. a region where feature and background have a constant grey level difference) there is always the difficulty of knowing which threshold level to set in order to get optimum feature extraction.

The first problem can be solved by splitting the image into a discreet number of smaller tiles in such a way that if a tile contains a feature, then the difference in grey level between the feature and the background is constant. If a tile contains multiple features, it could be further split into smaller tiles, until there is a maximum of one feature and one background in each of these smaller tiles. When the image has been split into the required number of tiles, each tile can be thresholded individually with a different level. Although this method of tiling may not be particularly robust, it is an improvement over trying to threshold a large image with just one threshold level.

The second problem, actually deciding which is the most appropriate level for thresholding, is more difficult to solve, and is compounded by the fact that if the image has to be tiled, it has to be solved many times over. One possible solution is to manually

try many different threshold levels on each image or tile until the required result is obtained, but this is very labour intensive. A second possibility is to automatically select the threshold level based on the content of the image or tile; this is known as automatic thresholding. A survey of automatic thresholding techniques is given by Weszka (1978). Since manual thresholding can be very time consuming, especially with large images, a technique of automatic thresholding has been developed in the course of this study.

Automatic thresholding uses the shape of the histogram to determine the most appropriate level for the threshold. Since the shape of the histogram depends on the content of the image, the success of automatic thresholding is also heavily dependent on the image content. For this reason, automatic thresholding is not appropriate for all types of images. The simplest type of image to threshold automatically is one which contains a single feature on a uniform background, since in this case the histogram will usually have two peaks, one representing the feature, and one representing the background. This type of histogram is known as a bimodal histogram. Figure 5.11 shows an example of this type of image, and its corresponding histogram.

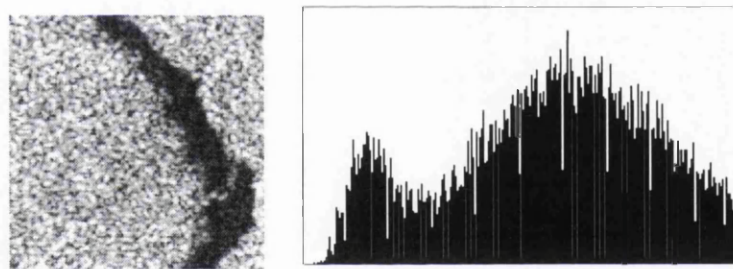


Figure 5.11 Single feature image (left) and corresponding bimodal histogram (right)

The image in figure 5.11 is a radar image of an oil slick at sea. In the histogram the dark oil slick is represented by the smaller peak on the left, whilst the background of the ocean is represented by the larger peak on the right. The histogram is very noisy because the radar image has not been smoothed in any way and therefore the speckle in the image is affecting the histogram. In a bimodal histogram the most appropriate level for the threshold is somewhere between the two peaks. From a manual examination of the histogram above, it was estimated that a threshold level of 100 would fall approximately half way between the peaks of the histogram. Figure 5.12 shows the result of thresholding the image at this level.

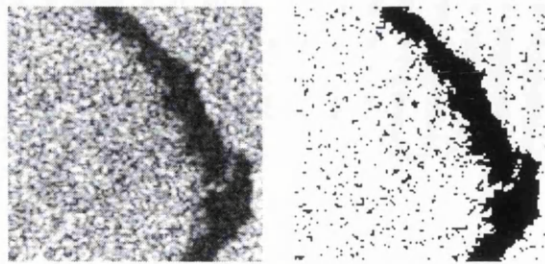


Figure 5.12 Radar image of oil slick (left) and image thresholded at grey level of 100 (right)

It can be seen therefore that by thresholding the image at a value which lies between the peaks of the bimodal histogram, the main feature can be separated from the background. A lot of clutter has also been extracted, but this can be simply removed by further processing (see § 5.4).

Therefore, to automate the extraction of features from images using thresholding the procedure must automatically determine the threshold level. This would be a trivial problem if the histogram had two clear peaks separated by a distinct trough. However, in general histograms are very noisy and contain many peaks and troughs superimposed on the bimodal shape, as can be seen in the histogram in figure 5.11. It is therefore necessary to process the histogram so that the locations of the two main peaks and the trough can be determined accurately. This can be done by smoothing the histogram, and a technique has been developed here which uses different smoothing kernels, and a varying number of iterations.

To smooth the histogram, the frequency of a particular grey value f_{DN} is averaged with its neighbours to give a new value of the frequency f'_{DN} at that grey level. This can be expressed by the equation:

$$f'_{DN} = \frac{1}{2n+1} \sum_{i=1}^n f_{DN-i} + \dots + f_{DN} + \dots + f_{DN+i} \quad (5.20)$$

For a 3 grey level kernel the value of n is 1, for a 5 grey level kernel n is 2 and for a 7 grey level kernel n is 3. Each of these three kernels were applied to the histogram of the oil slick image, and the resulting histograms are shown in figure 5.13.

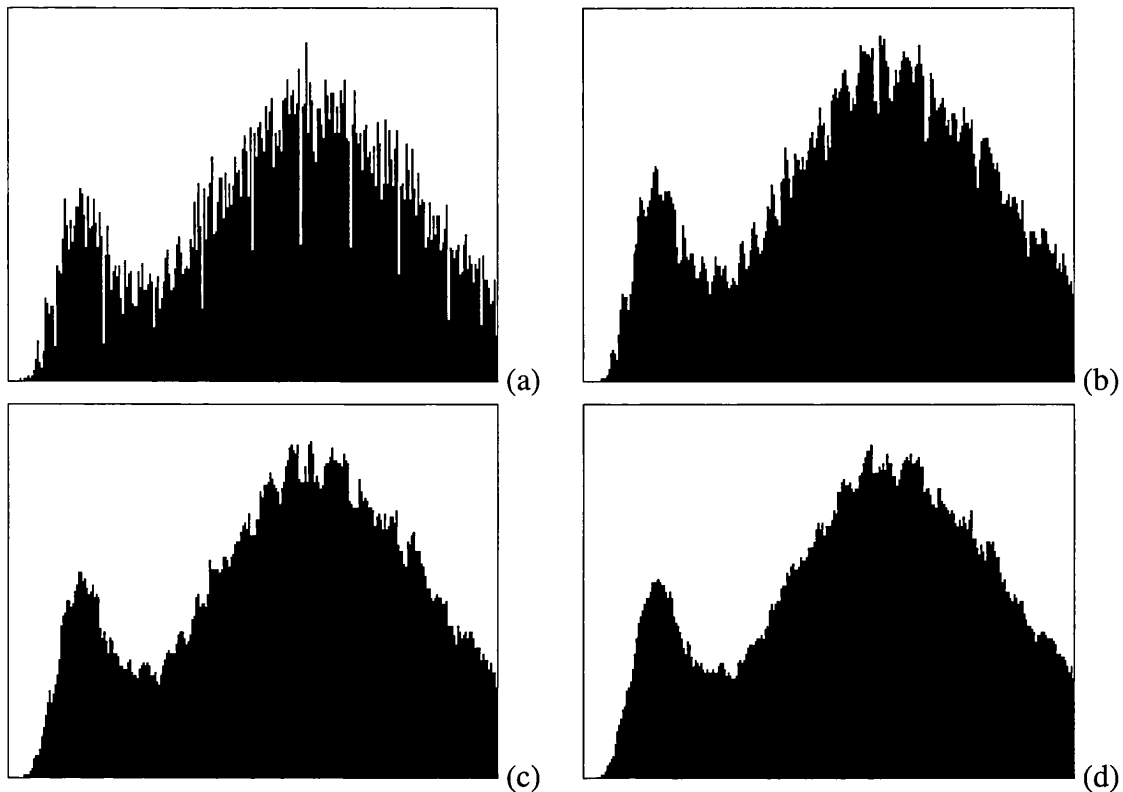


Figure 5.13 Original histogram (a), and histograms smoothed with 3, 5, 7 grey level kernels (b, c, d respectively).

It can be seen that although a larger kernel gives better smoothing of the histogram, even the largest kernel used here (7 grey levels) does not give a histogram with exactly two peaks. Therefore, rather than applying larger and larger kernels, it was decided that it would be more efficient if the same kernel was applied repeatedly to the same histogram. Figure 5.14 shows 10 iterations of the 7 grey level kernel to the oil slick histogram. The original histogram is shown as a red line, the 10th iteration is shown as a blue line, and all iterations between are shown as black lines. (Note that in the diagram, successive iterations have been drawn displaced from the previous, so that the curves do not overlay each other.) It can be seen that by the sixth iteration, there are exactly two peaks in the histogram.

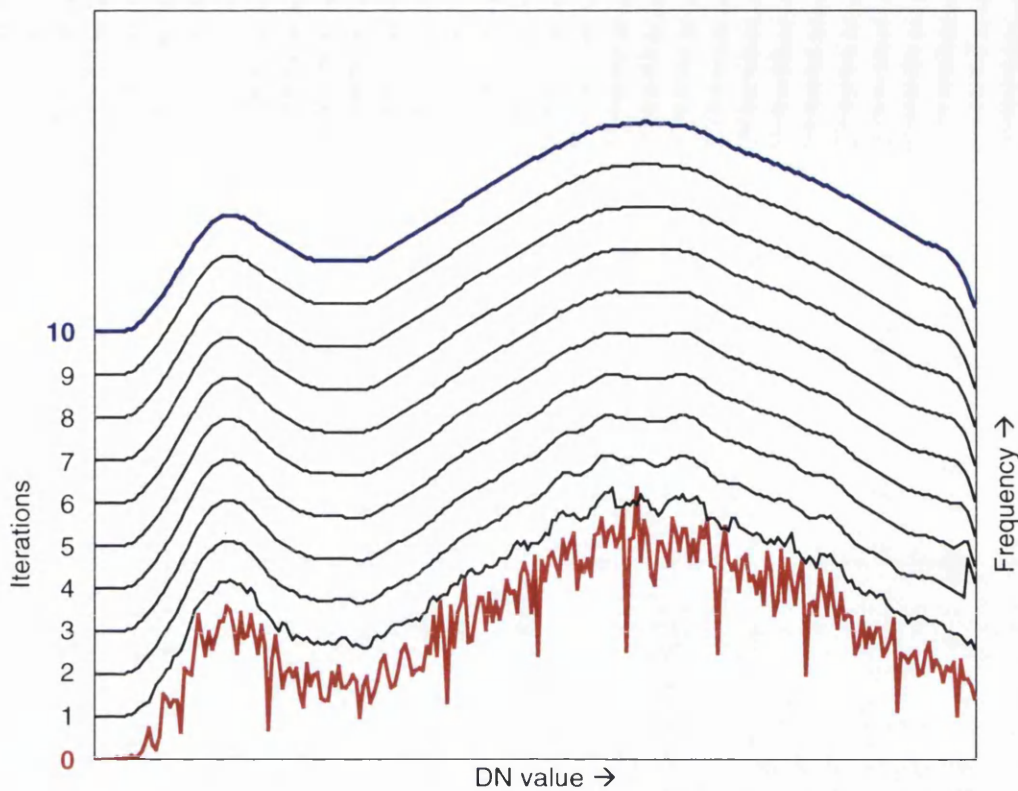


Figure 5.14 Ten iterations of histogram smoothing using 7 grey level kernel

The aim of histogram smoothing is to obtain a histogram with exactly two peaks. The correct combination of kernel size and number of iterations will achieve this, but this combination will be different for all images. For any image it is possible to construct a table giving the number of peaks in the histogram for different combinations of kernels and iterations. For the image in question, this table is shown below (table 5.1). Note that the table shows number of turning points, rather than number of peaks (these values the represent the same quantity since number of turning points = number of peaks + number of troughs).

		Number of iterations									
Kernel size		1	2	3	4	5	6	7	8	9	10
	3	98	61	49	45	31	31	21	19	19	17
	5	88	29	17	15	11	7	5	5	5	5
	7	82	21	7	5	5	3	3	3	3	3
	9	54	9	5	5	3	3	3	3	3	3
	11	54	7	3	3	3	3	3	3	3	3
	13	48	9	3	3	3	3	3	3	3	3
	15	40	5	3	3	3	3	3	3	3	3

Table 5.1 Number of turning points for different combinations of kernels and iterations

For each of the combinations of iterations and kernel size which give three turning points (in other words, two peaks and a trough) it is possible to calculate a threshold level. This threshold level can either be chosen as the location of the trough, or

somewhere else between the two peaks in the histogram. The trough of the histogram actually represents the pixels which lie on the boundary between the feature and the background. By thresholding at this level some boundary pixels will be classed as feature pixels and others will be classed as background pixels. By thresholding at a slightly higher level all of the boundary pixels will be classed as feature pixels. It was decided that this would give a better representation of the feature, so the threshold level was set at the level exactly half way between the trough and the background peak. This was done for the above example, and the threshold levels which were calculated are shown in table 5.2 below. Note that some of the cells in the table are empty since it is only possible to calculate a threshold level for three turning points.

		Number of iterations									
Kernel size		1	2	3	4	5	6	7	8	9	10
	3	-	-	-	-	-	-	-	-	-	-
	5	-	-	-	-	-	-	-	-	-	-
	7	-	-	-	-	-	94	98	98	98	98
	9	-	-	-	-	98	98	98	98	98	98
	11	-	-	95	97	97	98	98	98	99	99
	13	-	-	98	98	98	98	99	99	99	99
	15	-	-	98	98	99	99	99	100	100	100

Table 5.2 Threshold levels for different combinations of kernels and iterations

The threshold levels in table 5.2 above illustrate that whichever combination of iterations and kernel size giving three turning points is used to determine the threshold level, the results do not vary very much. In the above example, the mean, mode and median threshold level are 98.17, 98 and 98 respectively. This implies that in this example, the best threshold level to at which to threshold the image would be 98 (which actually agrees quite well with the initial estimate of 100). It was therefore decided that to automatically select the threshold level, the mean should be taken of all the threshold levels calculated from each combination of iterations and kernel size. To ensure that no spurious values occur, the mean should only be taken if there are a minimum of six combinations with three turning points.

This method of automatically determining the threshold level is now applied to two typical SAR images: an image featuring an oil slick at sea, and an image of agricultural land in Southern France. In the case of the oil slick, there is one object and one background, so the algorithm is able to calculate the most appropriate threshold level relatively simply since the histogram is obviously bimodal. The result shown in figure 5.15 clearly shows that the slick has been extracted well from the uniform background,

even though a lot of clutter has been extracted as well. (As already mentioned, this clutter can be easily removed using techniques described in § 5.4).

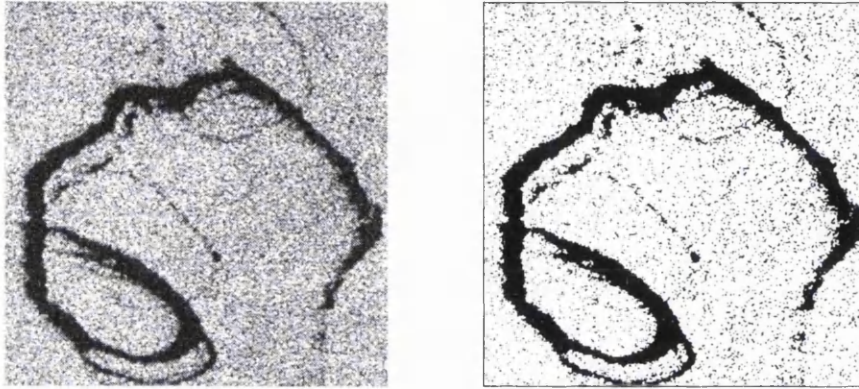


Figure 5.15 Oil slick image (left) and automatically thresholded image (right)

Unfortunately it is not always so simple to achieve a good result like this. Often an image gives the impression that its histogram will be bimodal when this is not the case. Figure 5.16 shows the second SAR test image, featuring a river passing through the town of Beaucaire in Southern France. On the right is the corresponding histogram.

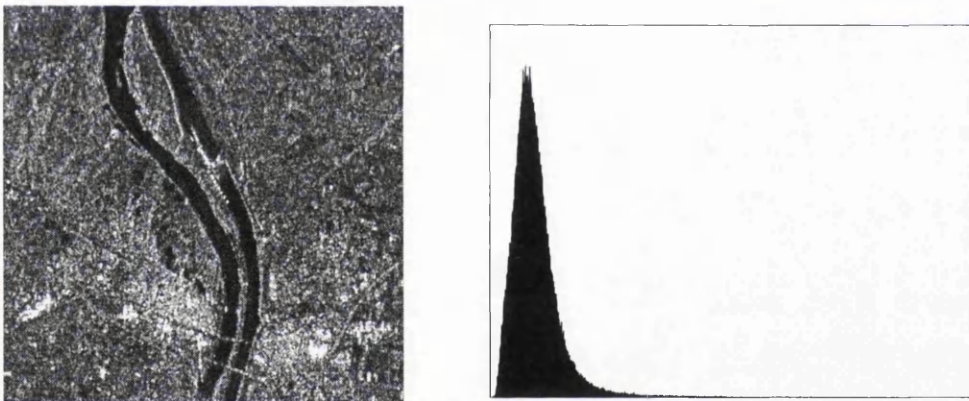


Figure 5.16 SAR image of river and town (left) and corresponding histogram (right)

It would be expected that the histogram would be bimodal, with a small low grey value peak representing the river, and a large higher grey value peak representing the background. Two possible reasons that this is not the case is that the feature is relatively small in comparison to the size of the image, and that the speckle is saturating the histogram and obliterating any small peaks. Although it is not possible to increase the size of the object in this image, an obvious way overcoming the problem of speckle corrupting the histogram is to simply remove the speckle from the image. By smoothing the image, more pixel values will be assigned to either the grey level of the object, or the grey level of the background; there will be fewer pixels with grey levels in between. This should lead to a more bimodal histogram. Figure 5.17 shows the histogram of the original image, and the histograms of the image after it has been smoothed with a mean

filter of increasing kernel sizes. (Note that a linear contrast stretch has been applied to the image to convert it from 16 bit to 8 bit; the maximum grey value in the 8 bit image (255) represents grey level 511 in the original image).

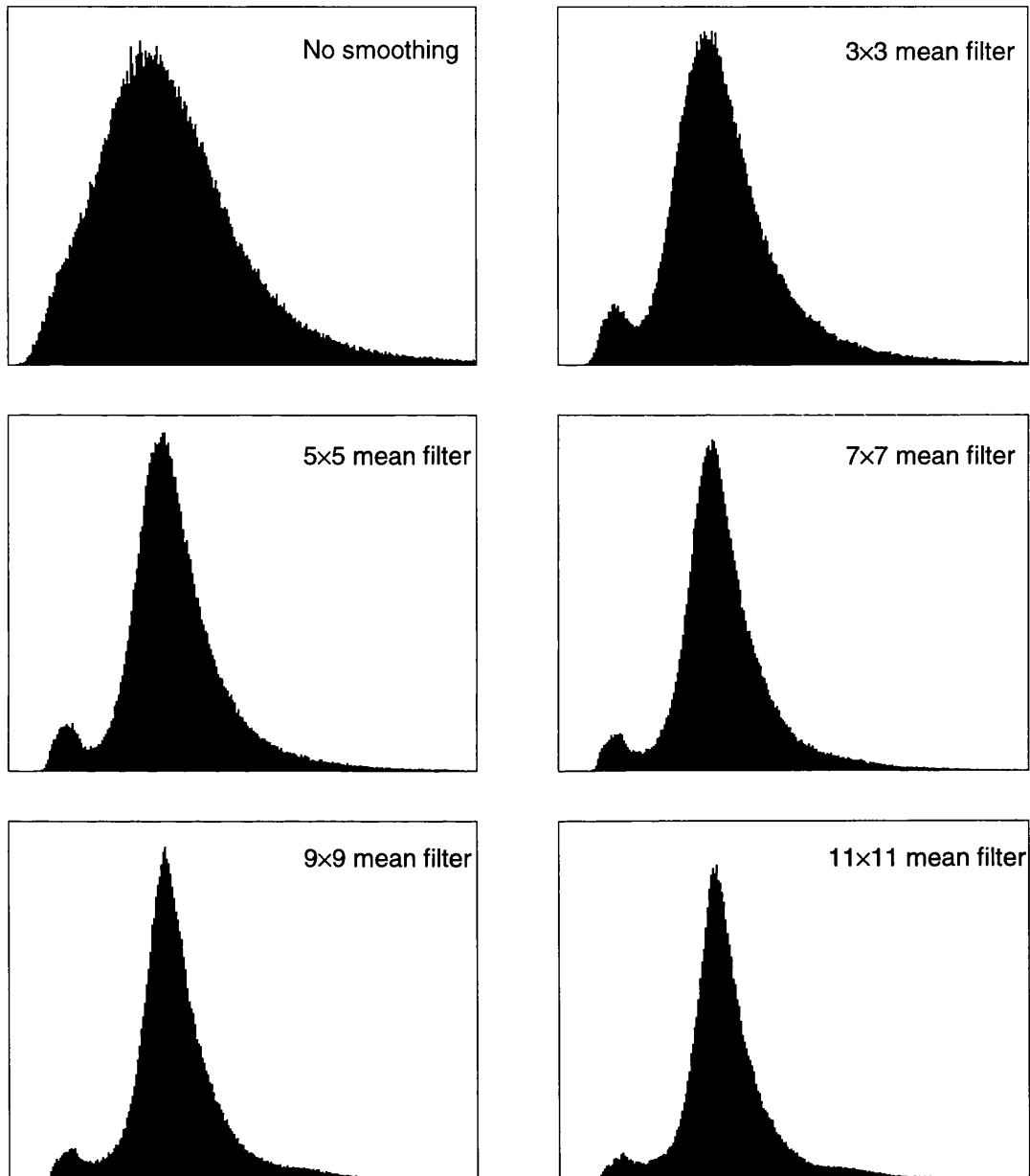


Figure 5.17 Histograms of smoothed Beaucaire image

The above figure shows that by applying smoothing filters to the image, the bimodal nature of the histogram can be accentuated. It appears that the best result is given by the 5×5 mean filter – larger filters lead to the peaks being smoothed too much. The automatically determined threshold value calculated from the image smoothed with the 5×5 mean filter is 60. The result of thresholding the original image at this level is shown in figure 5.18.

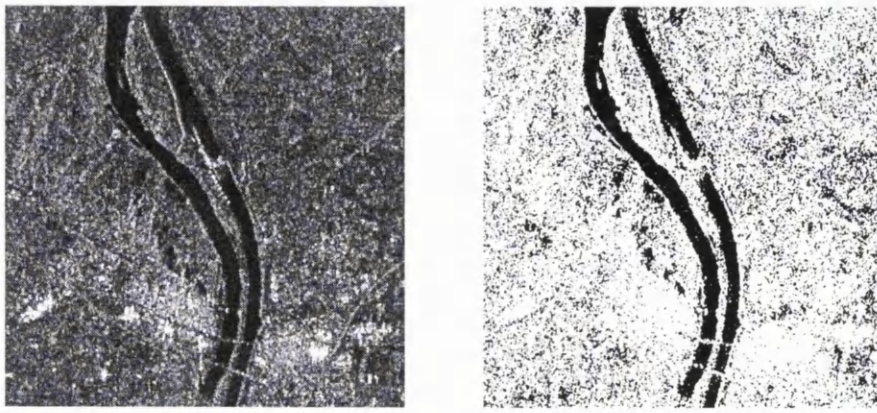


Figure 5.18 Automatically thresholded SAR image of region near Beaucaire

Figure 5.18 shows that by thresholding at the automatically determined value, some of the features in the image (the river and some dark fields) have been separated from the background. Once again, the clutter that has been extracted can easily be removed with post-processing.

The work so far has shown that it is possible to automatically extract major features from certain SAR images using thresholding. An attribute of all the images used so far is that they have all contained one major feature and a fairly continuous background. This is not true for all SAR images, and especially not true for optical images. The next part of this section deals with the extraction of features from SPOT images using automatic thresholding. Figure 5.19 shows a typical SPOT image of agricultural land in Southern France on the left, and on the right the histogram of this image.

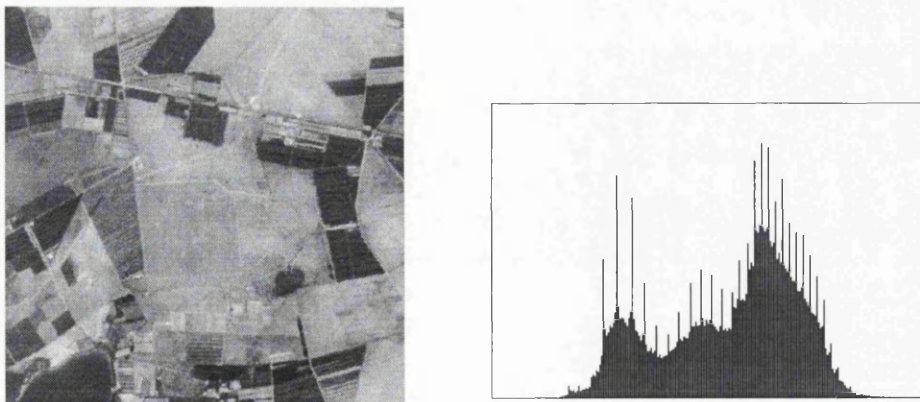


Figure 5.19 SPOT image of agricultural land near Entressen (left) and corresponding histogram (right)

It can clearly be seen from both the image and the histogram that there are more than one type of feature and background. When the automatic threshold algorithm described above is applied to this image, it determines the ideal grey level for thresholding to be 119. The result of thresholding the image at this level is shown in figure 5.20.



Figure 5.20 SPOT image of agricultural land thresholded at grey level 119

The dark fields have been separated from the lighter background, but some other features, such as lighter colour fields have been merged with the background. The problem here is that the histogram is multi-modal rather than bimodal, which is clear from figure 5.19. There is more than one major feature in the image, and there is a separate peak for each feature, so therefore the histogram has multiple peaks. It is not possible to split a multi-modal histogram at just one threshold level: instead a density slicing procedure needs to be used to split the histogram at each trough and thus separate each feature from the others. In this case the histogram smoothing operation has removed all but two of the peaks from the histogram and has therefore been able to determine a threshold level, which represents the optimum thresholding level for separating the major dark features from the lighter background. Automatic density slicing is much more difficult than automatic thresholding since the number of major features in the image must be known in order that the correct amount of histogram smoothing can take place to ensure that important peaks are not smoothed away. There is no obvious way of doing this without human intervention, so no attempt has been made here. However, it would make for a very interesting area of future research.

One of the drawbacks of using thresholding for the extraction of patches is that the patches which are extracted are not unique. They all have the same grey value, and are sometimes joined to each other by small groups of pixels. Therefore, it is not always possible to associate a particular patch extracted from the image with a single feature on the ground. When this is the case, it is more useful to convert the patch image into an edge image, and use the edges in the second stage of the image registration procedure, rather than using the patches in the first stage of the registration procedure.

In summary, this section has given a description of thresholding and shown how information contained in the image histogram can be used to automatically estimate the

ideal threshold level. A method of removing noise from the histogram has been presented which ensures that the correct peaks and troughs are identified. It has also been shown that smoothing the image can help to improve the bimodal nature of the histogram, which in turn helps to ensure that the correct peaks and troughs are identified. The algorithms described in this section have been illustrated using a selection of images that provide interesting results. Not all images provide useful results, and to show poor results would not add anything to this discussion. It is for this reason that the test images have been carefully selected, and null results have been left out.

5.3.3 Homogeneous patch extraction

Homogeneous patch extraction (Abbasi-Dezfouli and Freeman, 1994) is a very straightforward technique for extracting areal features from imagery. However, the drawback of the simplicity of this feature extraction method is that it will not consistently produce good results from all types of images. The algorithm works by scanning the image for homogeneous patches (i.e. regions with the same grey level) and separating them from the background. Since it is not common for features in the image to have a consistent grey value, a tolerance parameter can be set which specifies a range in which the grey values of pixels must fall.

The algorithm proceeds by firstly scanning the image for homogeneous 3×3 patches and extracting them. These 3×3 regions are then used as 'seed patches' from which larger homogeneous regions are grown. The process continues until none of the regions can be grown any further.

As the name suggests, only homogeneous patches can be extracted using this method. Features on the ground which are represented by continuous but textured regions in the image will not be extracted. In this study this should not prove to be a problem since the feature matching process relies on the results of a range of feature extraction algorithms, and not just this one. It is this redundancy that ensures the matching procedure produces consistent results.

Figure 5.21 shows a SPOT image of the region near Istres, and its corresponding features extracted using the homogeneous patch extraction algorithm. Figure 5.22 shows a SAR image of the same region and the corresponding extracted features. Tests showed that the results of patch extraction from the SAR images could be improved if 8 bit data was used rather than 16 bit data. Therefore a linear contrast stretch was performed on

the SAR images by assigning the values of 0 and 511 in the original histogram to 0 and 255 in the new histogram.

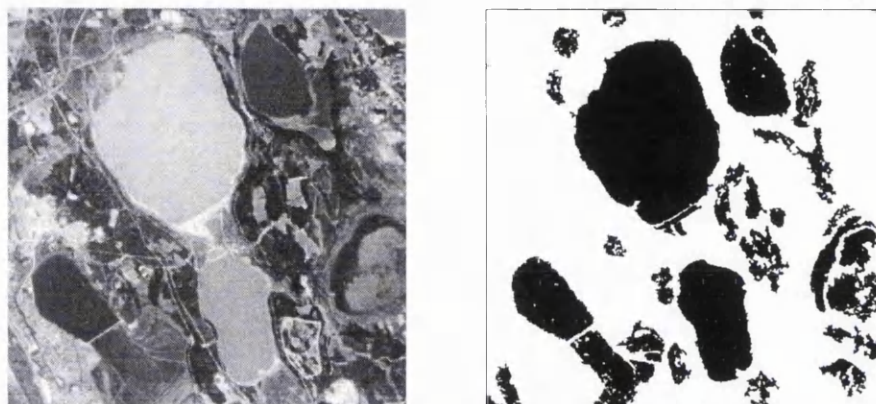


Figure 5.21 SPOT image of Istres (left) and extracted homogeneous patches (right)

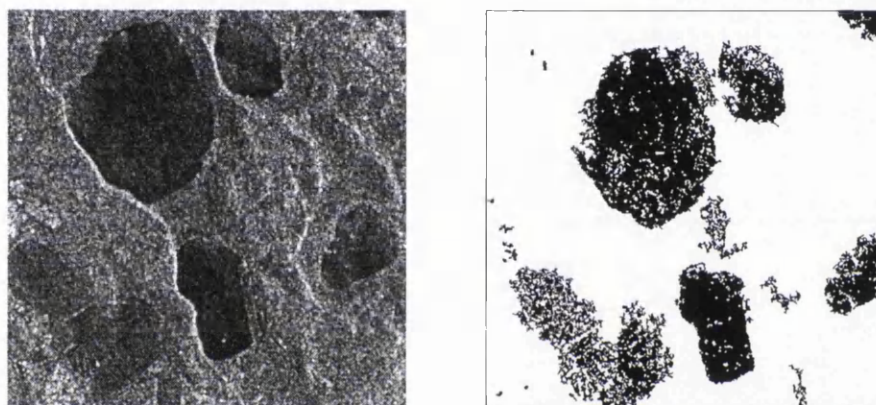


Figure 5.22 SAR image of Istres (left) and extracted homogeneous patches (right)

Patches were extracted from the SPOT image using a tolerance value of 5, while patches were extracted from the SAR image using a tolerance value of 13. As would be expected, more homogeneous patches have been extracted from the SPOT data than from the SAR data. This is because the SAR image is corrupted by speckle, and therefore features are much less homogeneous. However, both of the above results can be improved by post-processing: small patches can be removed from the images, and small holes in the patches can be filled. It is this combination of feature extraction and post-processing which makes homogeneous patch extraction a useful method of feature extraction for image registration.

The homogeneous patch extraction algorithm was also tested with images that had been preprocessed using the smoothing and speckle reduction algorithms described in chapter 4. For the SPOT images it was found that preprocessing made the results much worse. The smoothing algorithms led to the grey level differences between features in the image to be reduced, causing the extracted patches to become merged with each other. For SAR images the result of using the homogeneous patch extraction algorithm in

conjunction with the speckle reduction algorithms introduced in chapter 4 was that fewer features could be successfully extracted. Smoothing algorithms caused features to become merged, as with the SPOT images, and adaptive non-rigorous algorithms led to false patches being extracted. The rigorous adaptive algorithm performed slightly better, resulting in extracted patches with slightly fewer holes, but the results were not good enough to further investigation. It was decided that in the case of homogeneous patch extraction, more could be achieved with post-processing of the extracted patches than preprocessing of the images.

5.3.4 Segmentation

Segmentation is a very well established method of extracting features from images. It is similar to classification in that it splits the image into a number of discreet regions, but unlike classification, there is no limit to the number of classes. Furthermore, the method of deciding which pixels to group into which regions is quite different in classification and segmentation. Since the properties of SAR and optical images are quite different, the algorithms used to segment them are also different. For this reason, the description of the SAR segmentation algorithms, and the SPOT segmentation algorithms are separated below.

SAR segmentation

The two SAR segmentation algorithms which were used in this study (MUM and RWSEG) are commercial products supplied by N. A. Software Ltd., UK (Caesar, 1996). Each is slightly different in the way they work, and produce slightly different results. The MUM (Merge Using Moments) algorithm works by initially extensively over-segmenting the image and then merging neighbouring regions to produce a coarser segmentation (Cook et al., 1994). The merging uses statistical properties (moments) of the regions to decide whether they represent the same underlying cross-section; if so, they are merged. The process is repeated until no further merging is possible. A tuning parameter, p , can be set by the user to have some control over the final image segmentation: if the probability that two neighbouring regions represent the same underlying cross-section is greater than 10^{-p} , then the regions are merged.

The RWSEG algorithm works by successively detecting edges in an image, and growing regions between them (White, 1991). The first step in the procedure is the application of an edge detector which adapts its shape to local conditions. A pixel is

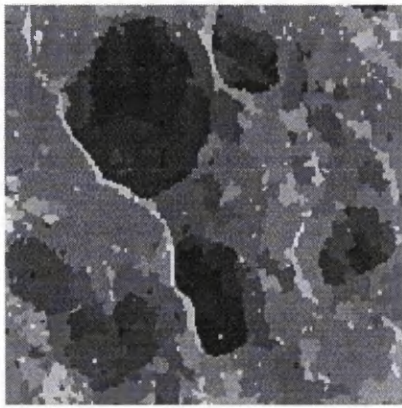
classified as an edge if the class separation distance between two regions is greater than some threshold, e . This can be expressed as:

$$\frac{|\mu_A - \mu_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}} \geq e \quad (5.21)$$

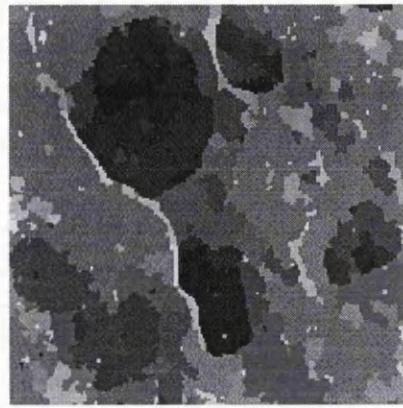
where μ_A and μ_B are the means of regions A and B, σ_A and σ_B are the standard deviations of the regions A and B, and e is a user defined. Assuming the mean values have a Gaussian distribution a value of e of 2.33 is equivalent to a 2% probability of false alarms. Lower values of e increase the probability of false alarms, meaning more false edges will be detected in the initial segmentation. After this initial edge detection, gaps between edges are filled with decreasing size discs which are merged with each other if they overlap and have similar means. When no more discs can be inserted or merged, the process repeats itself: the edge detection algorithm is applied again, and the regions between edges are filled. The process continues to repeat until the segmentation stops improving. The result is an image which has been over-segmented, so a final merging stage is used to eliminate weak edges. Once again, the class separation distance is used to find weak edges. If it is less than some threshold j (which is user defined), the edge is defined as being weak, and the neighbouring regions are merged.

Since both of these algorithms have been developed from a statistical model of speckle they must be applied to uncorrelated SAR images. Therefore it is not possible to preprocess the images beforehand in order to improve the results.

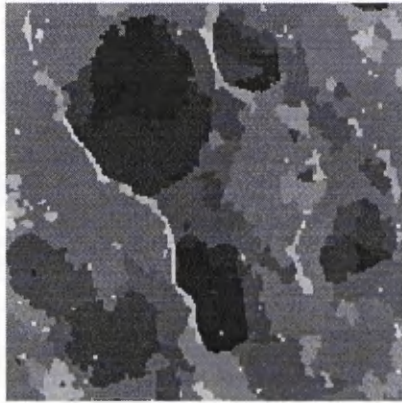
Figure 5.23 shows the MUM segmentation algorithm applied to the SAR image of Istres with four different values for the parameter p . Figure 5.24 shows the results of the RWSEG segmentation algorithm with two different values of the parameters e and j being applied to the same test image. With these segmentation algorithms it was possible to use the original 16 bit data rather than having to perform a linear contrast stretch.



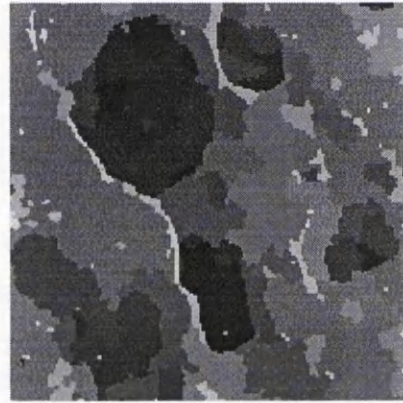
$(p=3)$



$(p=4)$

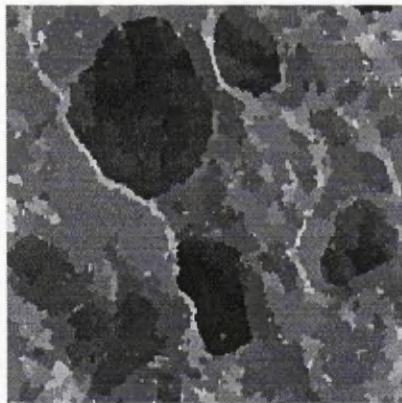


$(p=5)$

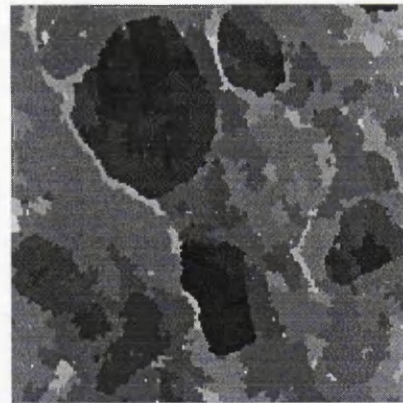


$(p=6)$

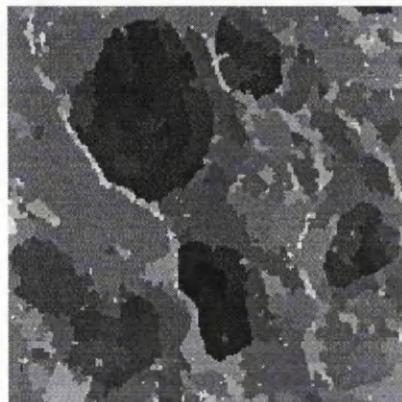
Figure 5.23 SAR image segmented with MUM using four different values of p .



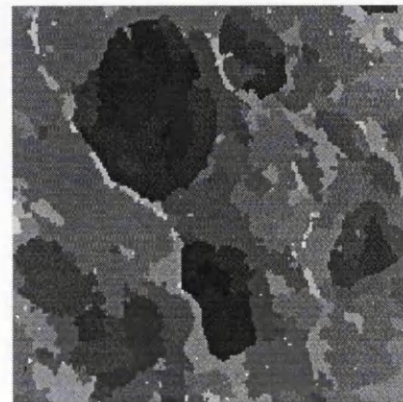
$(e=2.81, j=0)$



$(e=2.81, j=3.5)$



$(e=3.29, j=0)$



$(e=3.29, j=3.5)$

Figure 5.24 SAR image segmented with RWSEG using two different values of e and j .

The above figures show that both the MUM and RWSEG algorithms can segment the SAR images very successfully. It can be seen in figure 5.23 (the MUM segmentation) that increasing the value of p reduces the number of regions, but the effect is not great. The RWSEG algorithm, with the parameter j set to 0, can be seen to produce more regions in the segmented image than the MUM algorithm, but increasing j to 3.5 merges the regions to such an extent that the result is similar to that of the MUM algorithm. Increasing the value of e from 2.81 to 3.29 reduces the number of regions, but the change is not very great. It is difficult to compare the outputs of these different algorithms since there are so many different combinations of parameters that can be used. The four combinations used here were chosen to exemplify the type of results that can be obtained and how they differ. The question of how to select the optimum parameters for image segmentation is addressed in the next chapter.

Optical segmentation

Two segmentation algorithms used for extracting features from optical images which were tested in this thesis were REGSEG (Kai and Muller, 1991) and OPTISEG (Ruskoné and Dowman, 1997).

The REGSEG algorithm is a region growing segmentation process in which regions are grown one-by-one in all directions from seed points. Growing continues until an edge pixel is met, or the difference between a growing region and a potential candidate pixel for that region is above some threshold. This threshold, d , has to be set by the user before processing takes place. The advantage of this method of image segmentation over standard region growing techniques is that regions are grown in all directions from seed points, rather than being grown in a particular scanning direction, therefore making the algorithm more robust. Furthermore, the optional use of edge information allows a limit to be set on the size and shape of the regions. This makes the REGSEG algorithm very flexible since the user is not obliged to use edge information, and if edge information is used it does not have to come from the image being segmented – it could come from digital map data, or another image, as long as those data sets were registered to the image being segmented.

The OPTISEG algorithm is loosely based on the REGSEG algorithm, but is much more advanced and therefore tends to produce better results. It works in the same way as REGSEG: regions are grown based on clustering of neighbouring pixels which have similar properties. A cluster of pixels forms a region, and further pixels are added if

they fulfill a criterion based on the grey level difference between the pixel and the region, and the mean grey level of the region. This criterion can be expressed as a threshold, T , where:

$$T = b + \mu M_i - \nu V_i \quad (5.22)$$

where b is the difference between grey level of the candidate pixel and grey level of region i , M_i and V_i are the mean and variance of region i , and μ and ν are user defined variables. In order to segment an image, the user must select values for T , μ and ν as well as a further parameter which sets a lower limit for the size of regions. Once the segmentation process has finished an image is obtained which is over-segmented (as was the case with the RWSEG algorithm), so further processing has to take place to merge neighbouring regions together. The three types of regions which have been over-segmented, and therefore can be merged, are:

- regions split by local distortions;
- small regions wholly within larger region which were formed by noise pixels;
and
- regions in heterogeneous areas.

To merge regions that have been split by local distortions, the radiometric similarity of the regions is determined using the class separation distance. This value is compared with a threshold and regions are merged accordingly. The process repeats with a linearly decreasing threshold value until no further changes take place.

The ratio between the length of the longest boundary and the area of a region is used to decide whether that region is wholly enclosed within another region, and therefore whether the region should be merged with its surroundings. If the ratio is above some threshold, then the merging takes place.

The third type of over-segmentation is resolved by evaluating a shape measurement, g , for each region, where:

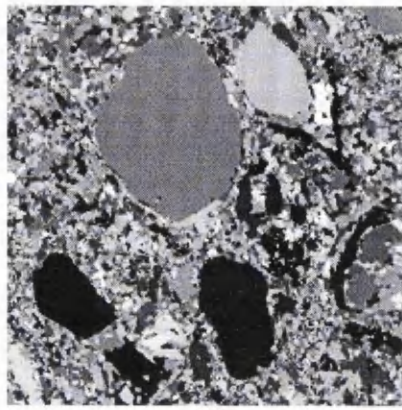
$$g = \frac{A}{d^2} \quad (5.23)$$

where A is the area of the region and d is the mean minimum distance of every point in the region to the boundary. Shapes that are very jagged have low values of A , and they can therefore be merged with their neighbours.

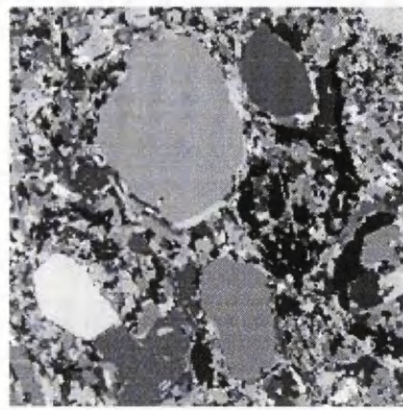
Note that without the post-segmentation merging of regions, and with μ and ν both set to zero, the OPTISEG algorithm will produce exactly the same results as the REGSEG algorithm.

A feature of both the REGSEG and OPTISEG algorithms is that they can be used with images that have been preprocessed. Both algorithms were tested with three different images: non-smoothed, smoothed using the median filter, and smoothed using the MHN filter. It was found that the best results were achieved when the MHN filter was used to preprocess the images. Without this filter the images tended to become under-segmented. Smoothing with the MHN filter meant that low values of the pixel difference parameters (d for REGSEG and b for OPTISEG) could be used without the risk of under-segmentation.

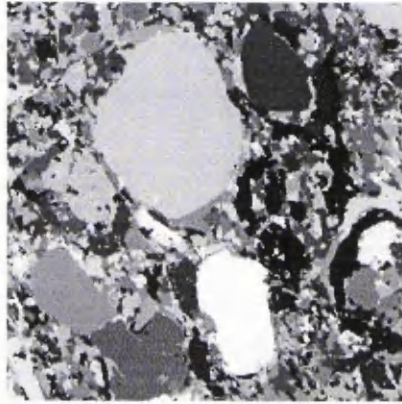
The two segmentation algorithms described above (REGSEG and OPTISEG) were applied to the SPOT image of Istres. Figure 5.25 shows the result of applying the REGSEG algorithm with four different values for the parameter d ($d = 4, 5, 6$ and 7). Figure 5.26 shows the result of applying the OPTISEG algorithm with various combinations of values for T , μ and ν . Note that the grey values of the output regions are random so the spectral properties of the segmented image does not reflect the spectral properties of the input image.



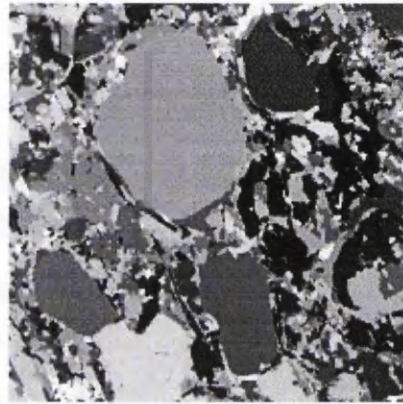
($d=2$)



($d=3$)

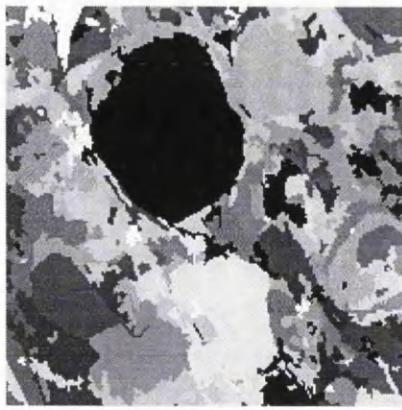


($d=4$)

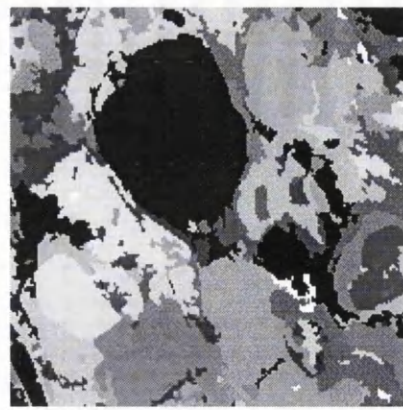


($d=5$)

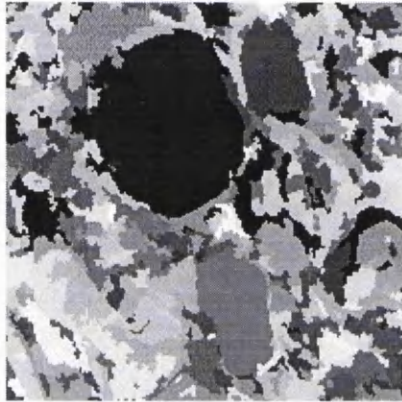
Figure 5.25 SPOT image segmented with REGSEG using four different values of d (2, 3, 4 and 5).



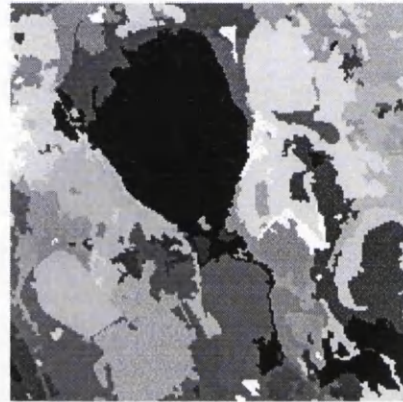
$(T=1, \mu=0.2, \nu=-0.2)$



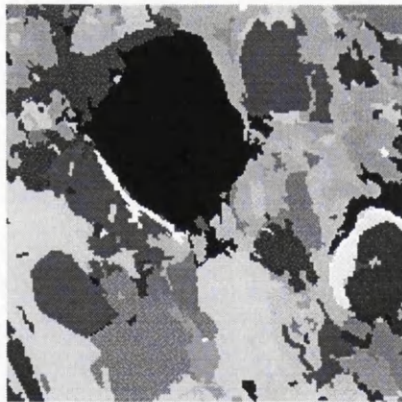
$(T=20, \mu=0.2, \nu=-0.2)$



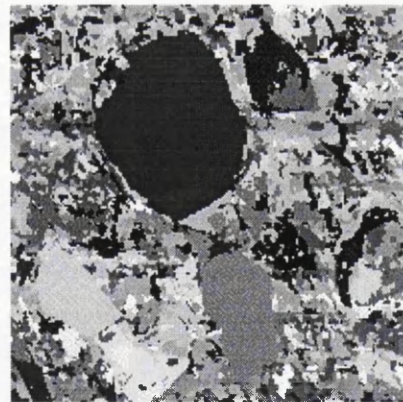
$(T=5, \mu=0.05, \nu=-0.2)$



$(T=5, \mu=0.5, \nu=-0.2)$



$(T=5, \mu=0.2, \nu=-0.05)$



$(T=5, \mu=0.2, \nu=-0.5)$

Figure 5.26 SPOT image segmented with OPTISEG using different values of T , μ and ν .

The images in figure 5.25 and 5.26 above show clearly that the two algorithms (REGSEG and OPTISEG) produce very different results. For low values of d , the REGSEG algorithm segments the large features well, but completely over-segments smaller features in the background. As the value of d is increased, the smaller features are not so over-segmented, but the larger features begin to merge with surrounding features. With $d=5$, it can clearly be seen that the image is under-segmented, i.e. too many features have been merged together to produce too few segments. Therefore, the most obvious aspect of the REGSEG algorithm is that the result is very finely dependent on the value of the parameter d .

The three pairs of images in figure 5.26 were generated by keeping two of the three parameters constant while adjusting the third. The first pair of images were generated using values of 1 and 20 for the parameter T . It is clear that by increasing T , the amount of segment merging increases and the number of segments decreases. Although large values of T lead to an under-segmented image, low values of T do not appear to over-segment the image. Increasing the parameter μ from 0.05 to 0.5 leads to a large change in the result. The lower the value of μ , the more segmented the image becomes. As with the parameter T , under-segmentation seems to be more typical than over-segmentation. The last parameter, ν , has the greatest effect on the result. By changing ν from -0.05 to -0.5 the image goes from being under-segmented (too few patches) to over-segmented (too many patches). Therefore the OPTISEG result is very finely dependent on the value of ν , just as the REGSEG result was finely dependent on the parameter d .

In conclusion, it can be said that both algorithms segment the image well, even though they give quite different results. The REGSEG algorithm will be useful for isolating large patches, whilst the OPTISEG algorithm will be useful for isolating both large and small patches. However, the larger number of parameters associated with the OPTISEG algorithm will make it slightly more difficult to fine tune when trying to produce the best results.

As mentioned above, both of these segmentation algorithms can be improved by incorporating an edge image into the processing, which can limit the size and shape of the output regions. This was tested using an edge image generated by locating zero crossings in the original image. The REGSEG segmentation of the Istres image was repeated incorporating this edge image (figure 5.27), and the results are shown in figure 5.28.



Figure 5.27 Edge map of SPOT image of Istres

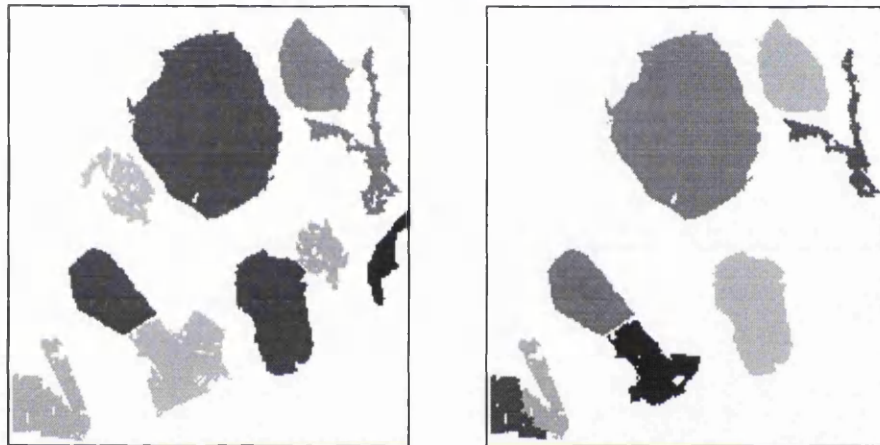


Figure 5.28 SPOT image segmented without edge map (left) and with edge map (right)

In the pair of images above, the one on the left has been segmented without the use of edge information, whilst the one on the right has been segmented with edge information. For clarity, only the largest segments have been displayed. When edge information is not used it can be seen that there is extra merging of regions (see, for example, the region in the bottom left hand corner of the images) which leads to there being more larger patches. With edge information, the growing of the patches is limited by the edges, so there tends to be smaller patches.

5.3.5 Boundary to region conversion

A final method of extracting areal features from images is by extracting boundaries of features (using techniques described in § 5.2), and using these boundaries to create polygons. Although it was only partially tested in this study, it is important to include a note about this process here since it completes the picture of linear and areal feature extraction. There are a few different methods of converting boundaries to regions, depending on the type of edge image used, and the type of result which is required. If the extracted boundaries form closed loops, the loop can simply be filled in order to

convert it into a region. If the boundaries do not form closed loops a procedure to extrapolate edges and connect them can be used to create closed loops, which can then be filled. Alternatively the areas between fragments of edges can be filled by inserting discs of decreasing sizes until regions are created with no gaps between them. This technique operates in a very similar way to the RWSEG algorithm, which segments an image by detecting edges and then creating a segmented image from that edge image. However, RWSEG repeats the process a number of times, and outputs a wholly segmented image, whereas boundary to region conversion repeats the process only once and is not limited to processing the entire image: as many or as few patches can be created as required. These different methods are described in more detail in Newton (1993) and Newton et al. (1994).

In this study an algorithm was developed for filling closed loops to create regions. Although this algorithm was successful, it was felt that developing methods of creating closed loops from open loops by edge interpolation would digress too far from the theme of the project. However, within the context of a robust system of feature extraction this would be an important tool, so although it is not developed further here, further research is warranted.

5.3.6 Conclusions

This section has introduced the subject of extraction of areal features from satellite imagery. Three methods of extraction have been investigated: thresholding, homogeneous patch extraction, and segmentation. One further method of detecting areal features by extracting boundaries and converting them to regions was described but only partially tested. The main reason for this is that the procedure may appear to be simple, but in reality it is actually quite complex, and to develop it fully would digress too far from the principle research themes of this study.

The technique of automatic thresholding described above has shown that for images with distinct features on a homogeneous background, it is easily possible to extract those features fully automatically using a thresholding technique. The advantages of this method of feature extraction is that it is simple and quick to perform, and does not require any user intervention at all. However, the drawbacks are that it is dependent on image content (meaning that it will not necessarily produce useful results for all images), and the features extracted can sometimes be difficult to separate from each other. This second drawback can be overcome by using the extracted features as edges

in the second stage of image registration, rather than patches in the first stage of image registration. Although this limits the usefulness of thresholding as method of feature extraction slightly, it is not a considerable problem.

One of the principle difficulties involved with automatic thresholding is processing the histogram so that it becomes truly bimodal (i.e. it only has two peaks and one trough). Due to noise, a histogram which has a bimodal shape may not be truly bimodal, making it quite difficult to select the exact locations of the major peaks and troughs. A method of processing using different size kernels and varying iterations was shown to give good smoothing and robust estimations of the locations of the major peaks and troughs.

The method of extracting features by identifying homogeneous patches proved to be surprisingly successful considering the simplicity of the method. Not so surprising was the fact that far better results were achieved with SPOT data than with SAR data, due to the poorer signal to noise ratio of the SAR images. Post-processing algorithms are able to improve the usefulness of the images for the matching procedure, making this method of feature extraction quite important to the automatic image registration procedure as a whole. However, it is still true to say that homogeneous patch extraction would benefit from further research.

The segmentation algorithms described in this chapter were all developed by other researchers, so the only work that could be carried out here was a comparison of the algorithms, and an investigation into how the algorithms could be tuned to optimize the extraction of features from SPOT and SAR data. Two SAR segmentation algorithms (MUM and RWSEG) and two SPOT segmentation algorithms (REGSEG and OPTISEG) were introduced. It was shown that MUM and RWSEG produce similar results, even when the segmentation parameters were varied. REGSEG and OPTISEG however produced quite different results. REGSEG tended to either over-segment the image or under-segment the image. It was clear that the algorithm is very finely dependent on the tuning parameter, and optimal selection of this parameter is very difficult. OPTISEG on the other hand is much more stable. The three tuning parameters still have a marked effect on the output, but the final result is not quite so strongly dependent on the values of these parameters. Furthermore, the additional post-segmentation merging aspect of the OPTISEG algorithm gives improved results over the REGSEG algorithm since it eliminates the problem of over-segmentation. A final point to consider regarding the optical image segmentation algorithms is the use of edge

information to improve the result. It was briefly shown that edge information can lead to segments which more truly represent real features on the ground. This technique of improving the segmentation will be shown to be very useful in optimizing the feature extraction and feature matching procedures.

There are many other methods of extracting areal features from remotely sensed imagery, but no further techniques will be described here. The remainder of this chapter concentrates on the post-processing of the extracted features, both areal and linear, in order to improve their use in the feature matching procedure.

5.4 Post-processing of extracted polygons

5.4.1 Introduction

The two previous sections have described how linear and areal features can be extracted from remotely sensed images. As was mentioned in the introduction, not all the extracted features need to be used in the registration process. Therefore, before trying match features from different images, post-processing algorithms can be applied to the extracted features to eliminate those that appear not to be very useful. There are other forms of post-processing which are also necessary to improve the matching procedure, such as unique grey value assignment and polygon to edge conversion. Each of these topics are discussed in the next three sections.

5.4.2 Unique grey value assignment

Since the type of output image generated by the different areal feature extraction algorithms described in § 5.3 above can vary, it is necessary to process the output images so they are all in the same format. The output images differ as follows:

- ❑ Thresholding produces single bit binary images, so all the polygons have the same grey level (either black on a white background, or white on a black background).
- ❑ Homogeneous patch extraction produces the same output as thresholding: single bit binary images.
- ❑ Polygons produced using the RWSEG and MUM algorithms have grey levels which are equal to the mean of all the pixels in the original that make up that region.
- ❑ OPTISEG produces polygons with unique grey levels (i.e. every polygon has an individual grey level), but these grey levels bear no relation to the original image.

- REGSEG produces polygons with non-unique grey levels and which bear no relation to the original image.

The output required for the matching procedure is one where each patch in the image has a unique grey value (so that each patch can be uniquely identified), but these grey values do not have to be related to the grey values of the pixels which make up the features in the original image (since features imaged by different sensors cannot easily be matched using spectral attributes). Therefore, the ideal output is provided by the OPTISEG algorithm. In order to convert the outputs from the other algorithms to the same format, a simplified version of the OPTISEG algorithm was used. This algorithm, called ASSIGN_DN, uses the OPTISEG algorithm to segment the processed images, but with the parameters T , μ and ν all set to zero, and with the post-segmentation region merging procedure removed. The result is that each polygon in the processed image is given a new and unique grey value, but its size and shape is not changed in any way.

A further property of the output is that the lowest grey level value is zero, and the highest is equal to the number of polygons minus one. This is very useful for reducing the amount of memory each image occupies. For an image with up to 255 different grey levels, the most efficient way to store each pixel value is as an unsigned 8 bit integer. In computational terms, this means that each pixel will require one byte of memory space. However, if the greatest grey level in an image is greater than 255, but less than 65535, then the most efficient way to store each pixel is as an unsigned 16 bit integer, but this takes up twice as much memory space as the 8 bit integers. If the greatest grey level in an image is greater than 65535, then the pixels must be stored as a 32 bit unsigned integers. The advantage of having an image with polygons whose grey values start at zero and increase in steps of unity is that it is possible to store the image in its most (memory) efficient format. Not only does this save computer memory space, but it also means that algorithms take less time to process the images.

5.4.3 Clutter removal

Some of the outputs from the areal feature extraction algorithms described above not only contain the major features in the image, but also a lot of insignificant minor features (see figure 5.29). These small features, referred to here as clutter, can be removed from the images without jeopardizing the accuracy or efficiency of the matching algorithm; in fact, removal of clutter actually increases the efficiency of the matching procedure. Two different methods to do this were implemented in this study.

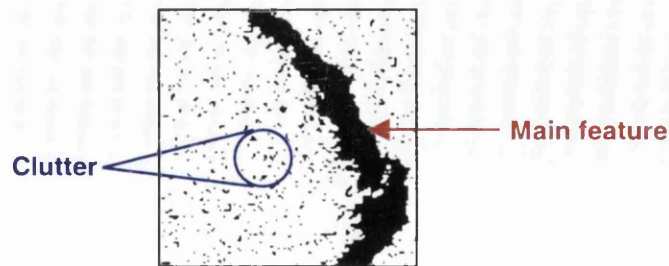


Figure 5.29 Clutter in automatically thresholded image of an oil slick

The first method of clutter removal (called RM_CLUT) scans the image with a square window and examines the pixels that make up the border of the window (i.e. the outside perimeter of pixels). If all the pixels in the border have the same grey value, then all the pixels within that border are set to the same grey value as the border. Figure 5.30 shows examples of 5×5 and a 7×7 windows.

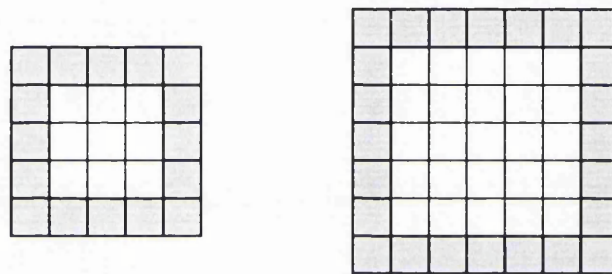


Figure 5.30 Examples of 5×5 and a 7×7 windows used for clutter removal

In the 5×5 window in figure 5.30, if all the pixels in the border (the grey pixels) have the same grey value, then all nine pixels within that border are set to the same grey level to the border, irrespective of what their values were before. Likewise, in the 7×7 window, if all the pixels in the border have the same grey value, then all 25 pixels within the border are set to the same grey level as the border. The algorithm is applied to the image by the user predefining a maximum window size and the filtering being repeated with increasing window sizes from 3×3 to that maximum size.

The second method of moving clutter (called RM_SMALL_PATCHES) is slightly more straightforward. Since every polygon has its own unique identity it is possible to filter polygons according to size. The user has to predefine the minimum allowable polygon size, and all polygons below this size are removed.

Both of these clutter removal algorithms work in different ways, and produce slightly different results, but both are very useful tools. RM_CLUT takes into consideration the shape of the polygon in question as well as the size, whereas RM_SMALL_PATCHES does not. RM_CLUT will not remove long thin features, even though they only consist of few pixels, whereas RM_SMALL_PATCHES will remove these features. However,

in an image which has been completely segmented, RM_CLUT will only remove polygons which are completely enclosed within other polygons, whereas RM_SMALL_PATCHES will remove all polygons below a certain size, irrespective of their surroundings in the image. Figure 5.31 shows the application of these two algorithms to the thresholded oil slick image.

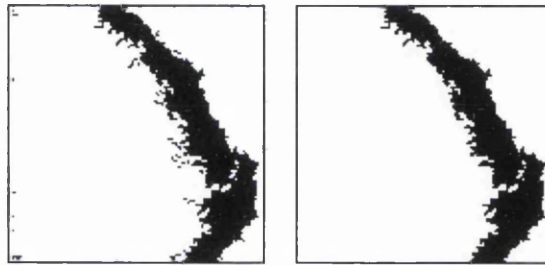


Figure 5.31 Thresholded oil slick image with clutter removed using RM_CLUT (left) and RM_SMALL_PATCHES (right).

The above images show that both of these algorithms perform very similarly, although RM_SMALL_PATCHES removes more clutter than RM_CLUT, especially in the area around the main feature. However, RM_CLUT is far more computationally efficient, and therefore is a useful tool.

5.4.4 Polygon to edge conversion

The final post-processing procedure described in this chapter is the conversion of polygons to edges. Linear features can be generated from extracted areal features by detecting the perimeter of the region and separating it from the rest of the image. It is not possible to apply an edge detection algorithm, such as those described in § 5.2 above, to the polygon image since the strength and direction values would be meaningless if the grey values of the polygons did not reflect the grey values in the image. Therefore an edge extraction algorithm was used which just located the edges between regions, and did not measure the strength or direction.

The edge extraction algorithm works by passing a 3×3 plus shaped kernel across the segmented image. If the central pixel has a different grey level to any of the other pixels in the kernel, then an edge is present (figure 5.32).

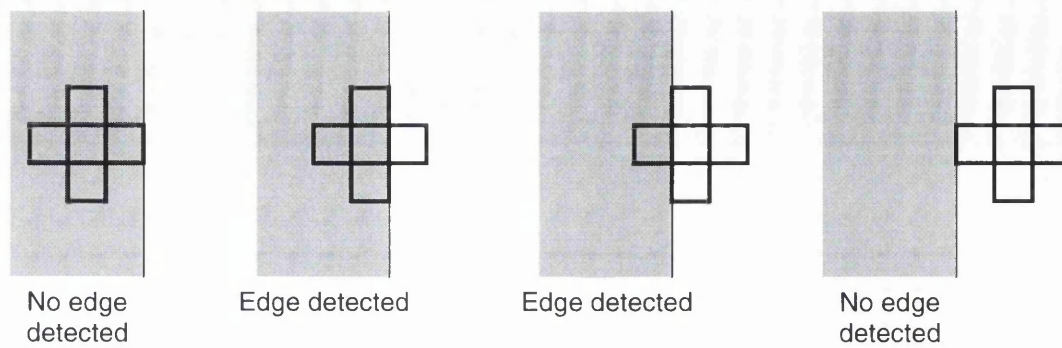


Figure 5.32 Four possible positions of edge extraction kernel in relation to an edge

Figure 5.32 shows four possible positions of the edge extraction kernel in relation to a boundary between two regions. In the first and last position no edge is detected because the central pixel has the same grey value as all the other pixels, but in the other two positions, edges are detected. The output is written to another image which becomes an edge representation of the segmented image.

5.4.5 Conclusions

This section has described some different post-processing techniques which can be used to enhance features, both linear and areal, extracted from images in order to improve the results of the feature matching process. These post-processing techniques may seem trivial in themselves, but are very important in the overall feature extraction and matching process. In the next chapters further processing techniques will be described which filter the images even further to improve the matching results.

5.5 Discussion

This chapter has shown that with some relatively simple feature extraction algorithms it is possible to perform some quite complex processing tasks on remotely sensed imagery and detect specific features. The purpose of this chapter was not to give a complete review of the latest methods of feature extraction, but rather to show how certain types of features can be extracted from particular imagery for the purpose of feature based image registration. The techniques described can generally be classed as low and medium level processing algorithms; high level processing has been avoided due to its complexity. The algorithms described here are tailored specifically to extract the most useful features for the method of image registration proposed in this study. Therefore they concentrate on extracting patches for the first stage of the registration procedure and edges for the second stage. The algorithms discussed in this chapter are summarized in table 5.3 below.

Algorithm	Image	Feature	Parameters
First derivative edge detector	SAR / SPOT	Edge strength and direction	Kernel size, threshold level
Second derivative edge detector	SAR / SPOT	Edge location	Filter width
Hough transform	SAR / SPOT	Linear feature	Line strength threshold
SCANEDGE	SAR	Edge location	Probability of false edges
Thresholding	SAR / SPOT	Regions and edge locations	Threshold level
Homogeneous patch extraction	SAR / SPOT	Regions and edge locations	Maximum grey level difference
MUM	SAR	Regions and edge locations	Probability of false patches
RWSEG	SAR	Regions and edge locations	Edge strength and merging parameters
REGSEG	SPOT	Regions and edge locations	Maximum grey level difference
OPTISEG	SPOT	Regions and edge locations	Maximum grey level difference, mean and variance parameters

Table 5.3 Summary of feature extraction algorithms

After features have been extracted, post-processing improves the quality of the result. The post-processing algorithms described in § 5.4 assign individual grey levels to each of the patches and successfully remove a lot of the clutter from the images, making further interpretation much simpler. The full usefulness of these two procedures will become clear in the following chapters when extracted features are matched with each other. Finally, a description of a method for converting patches to edges has been described. This is a very useful tool since it means that patches which are used in the first stage of the registration procedure can be converted to edges which can be used in the second stage. In this way the scope of some of the feature extraction algorithms is doubled.

A common drawback to almost all of the feature extraction algorithms described here is that they require some user intervention in the form of setting parameters. It has been

shown repeatedly in this chapter that optimizing the feature extraction algorithms by adjusting the parameters can lead to much better results. However, to fit in with the concept of an automatic registration system, it is necessary that the feature extraction stage of the procedure is also automatic. Therefore, this optimization of parameters must be performed automatically. The method of doing this is described in the next chapter.

An important improvement to image segmentation was suggested in § 5.3.4 where edge information from a second derivative edge detector was used in combination with an optical segmentation algorithm. The results showed that by combining edge information with the segmentation procedure lead to much improved results. However, the edge information can be generated by any of the feature extraction algorithms. Furthermore, multiple feature extraction algorithms could be used to generate the edge information. For example, an edge image could be generated by combining the results of a first derivative edge detector with linear features from the Hough transform and boundaries from thresholding. When this edge image is used in the segmentation procedure the results should be much improved.

As a final note it should be reiterated that the feature extraction algorithms used in the proposed method of automatic registration do not have to be able to extract all the significant features from an image. Multiple feature extraction algorithms will be used together, so those features missed by one algorithm will be extracted by one of the other algorithms. Secondly, it is not essential that all features are extracted from each image since only a few features are needed for the first stage of the registration. These two facts mean that the whole feature extraction procedure is much simplified.

Chapter 6 PATCH MATCHING

6.1 Introduction

The system of automatic image registration proposed in this thesis is a two step procedure. This chapter describes the first of these steps: patch matching. The justification for using patches in the first step of the matching procedure has been given in the previous chapters, as has a description of the methods used to extract patches from remotely sensed images. This chapter therefore concentrates solely the results of applying areal feature extraction algorithms to SAR and SPOT data, the matching of the patches, and the new techniques developed here to improve the quality of the matching results.

The first part of this chapter describes how the three methods of areal feature extraction (thresholding, homogeneous patch extraction and segmentation) were applied to the SAR and SPOT test images. Details of all the processing steps are given, including preprocessing (smoothing) and post-processing (clutter removal), and the results for each combination of image and algorithm is given. The subsequent section describes how the extracted patches are matched, based on attributes such as size, shape and location in the image. It also presents the results of the registration of the images based on the patch matching results. The last section describes new techniques for improving the matching results.

For the purposes of this part of the study two test areas have been selected to test the patch matching techniques: Istres and Camargue. These images were chosen since they display many different types of land cover and some interesting features with a range of sizes and shapes. The two pairs of images are shown in figures 3.20 and 3.22. The geometry of each pair of images above is very similar since the images were approximately aligned, the reasons for which were presented in chapter 2. To recap, approximate alignment is performed at the very beginning of the processing chain. It can either be carried out manually, by selecting a few tie points across the whole full

scene images, or automatically using ephemeris data supplied with the raw images. In this case the images have been approximately aligned using the manual method.

6.2 Extraction of patches from SAR and SPOT images

6.2.1 Introduction

This section shows the results of applying the areal feature extraction algorithms, described in the previous chapter, to both sets of SAR and SPOT test images. In each case a full description of how the algorithms have been applied has been given. Furthermore, each algorithm has been fine tuned to give the best possible result for the matching procedure. Details of this fine tuning has also been included.

6.2.2 Thresholding

Before any processing took place, the SAR images were converted from 16 bit to 8 bit by assigning the grey value of 511 in the input image to the value 255 in the output image. The histograms of the raw SPOT images were normalized to improve the contrast of the image. In the SAR Istres image there is one major object feature (water bodies) on a fairly continuous background so the automatic thresholding procedure can be applied. The histogram of the raw image was not significantly bimodal so it had to be enhanced using image smoothing. The result of automatically locating the peaks and troughs in the histogram lead to a suggested threshold level of 81. Figure 6.1 shows the original and thresholded images. In the SPOT image of Istres it is a bit more difficult to anticipate in advance whether or not the histogram needs enhancing since it is not clear whether there is only one object feature in the image. In fact, it turned out that no histogram enhancement was necessary and the automatic threshold routine determined the best threshold level to be 156. The results of automatically thresholding the SPOT image is shown in figure 6.2.

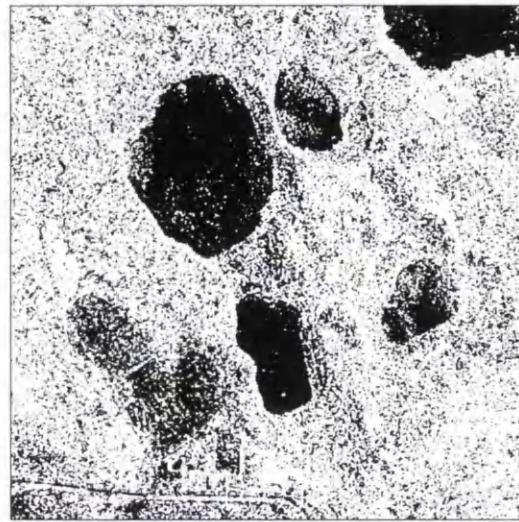


Figure 6.1 SAR image of Istres and corresponding thresholded image (threshold = 81)

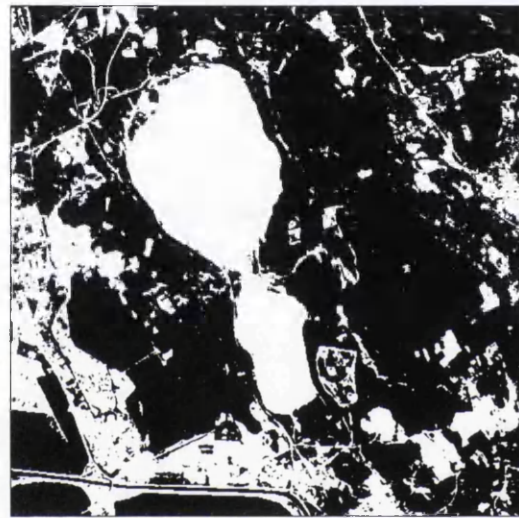


Figure 6.2 SPOT image of Istres and corresponding thresholded image (threshold = 156)

A similar procedure was followed for the pair of images of the Camargue, but both image histograms had to be enhanced by smoothing; the resulting threshold levels were 66 for SAR and 68 for SPOT. The original and thresholded images are shown in figure 6.3 and figure 6.4.

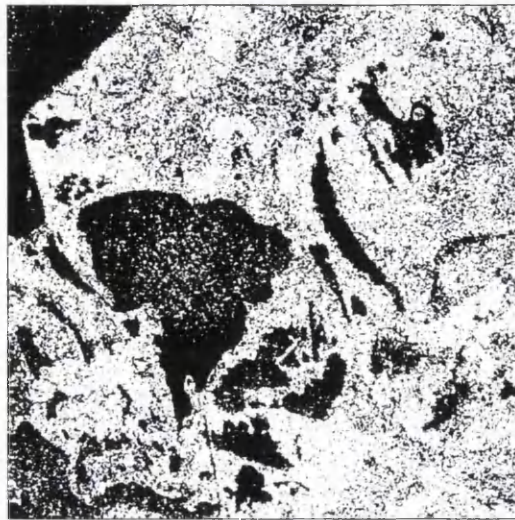
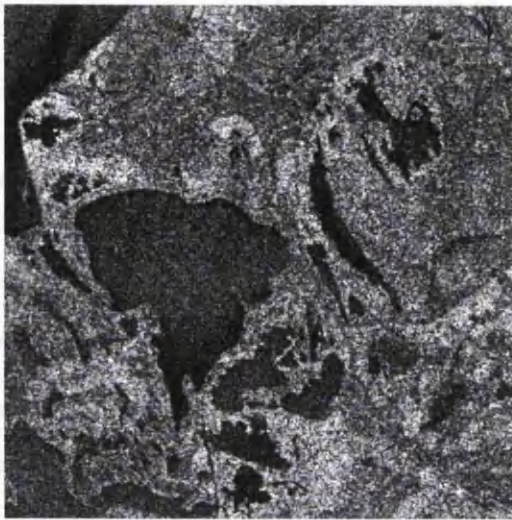


Figure 6.3 SAR image of Camargue and corresponding thresholded image (threshold = 66)

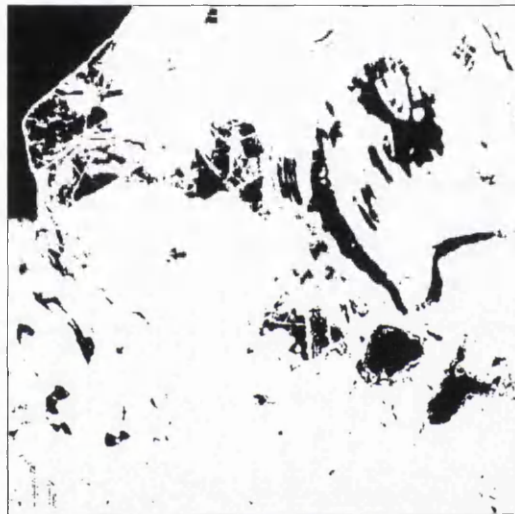
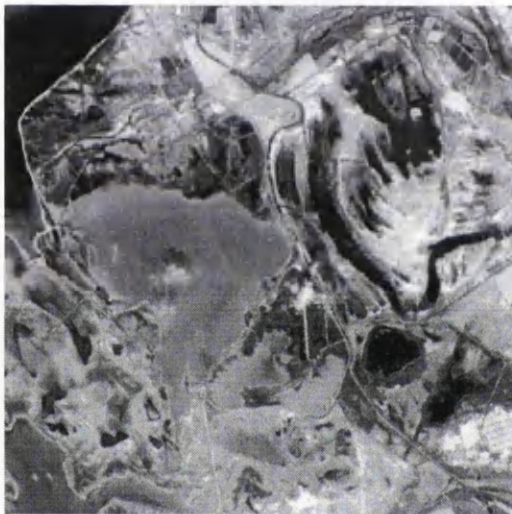


Figure 6.4 SPOT image of Camargue and corresponding thresholded image (threshold = 68)

As is common with thresholding, a lot of clutter has been extracted from the images along with the main features. It is not surprising therefore that a degree of post-processing clutter removal is required. This was done in two steps:

1. The algorithm RM_CLUT was used to remove clutter from both inside and outside the extracted regions.
2. The algorithm RM_SMALL_PATCHES was used to remove the remainder of the clutter missed by the RM_CLUT algorithm.

For the RM_CLUT algorithm the maximum window size was set to 15 for both images, and for the RM_SMALL_PATCHES algorithm the minimum patch size was set to 200 for both images. Both of these choices were fairly arbitrary, but seemed to give good results since most of the clutter was removed, but the important features were retained.

The results of the post-processing are shown in figure 6.5 and figure 6.6. (Patches in the SPOT image have been assigned random grey scale values for clarity.)

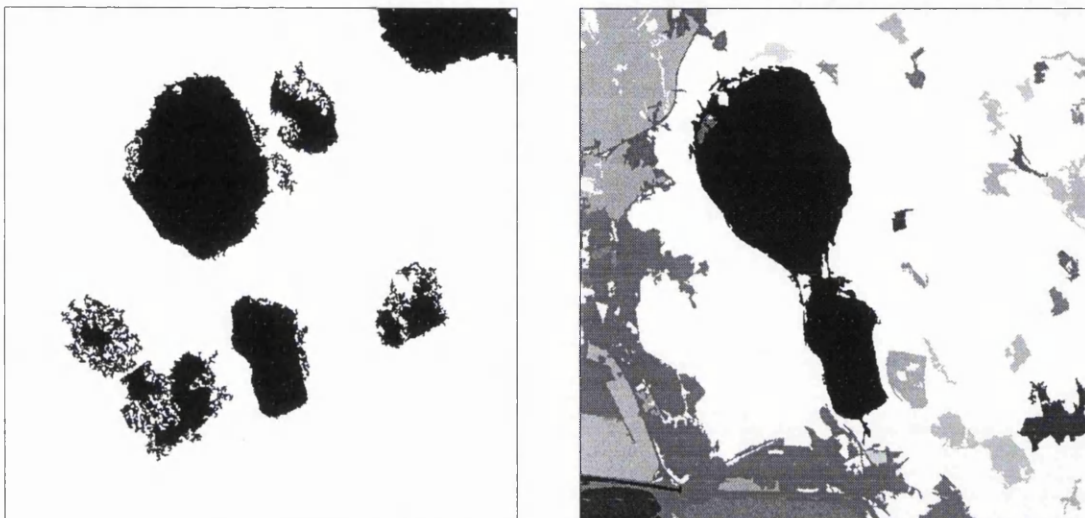


Figure 6.5 SAR and SPOT images of Istres after clutter removal

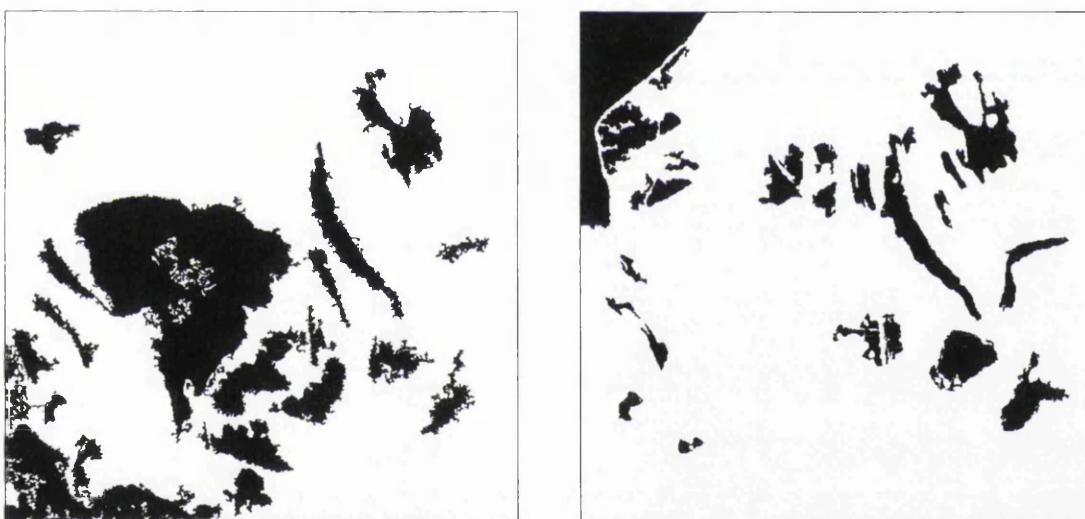


Figure 6.6 SAR and SPOT images of Camargue after clutter removal

The extraction of patches from these four test images using automatic thresholding has been relatively successful. In each case a threshold level could be determined automatically, although in three of the four images pre-smoothing had to be used to enhance the histograms in order to obtain a result. The features which have been extracted the best from the SAR images are the large water bodies, which is not really surprising since their grey values differ the most from the background. These water bodies have not been extracted quite so well from the SPOT images since the grey values differ much less. A lot of clutter has been extracted from the SAR images with the patches, but this was successfully removed using the clutter removal algorithms. The results show that a larger number of similar patches (i.e. those which represent the same features on the ground) have been extracted from the Camargue images than the Istres

images, leading to the conclusion that the Camargue images are better suited to matching based on patches extracted using automatic thresholding. A summary of the extraction of these patches using automatic thresholding is shown in figure 6.7.

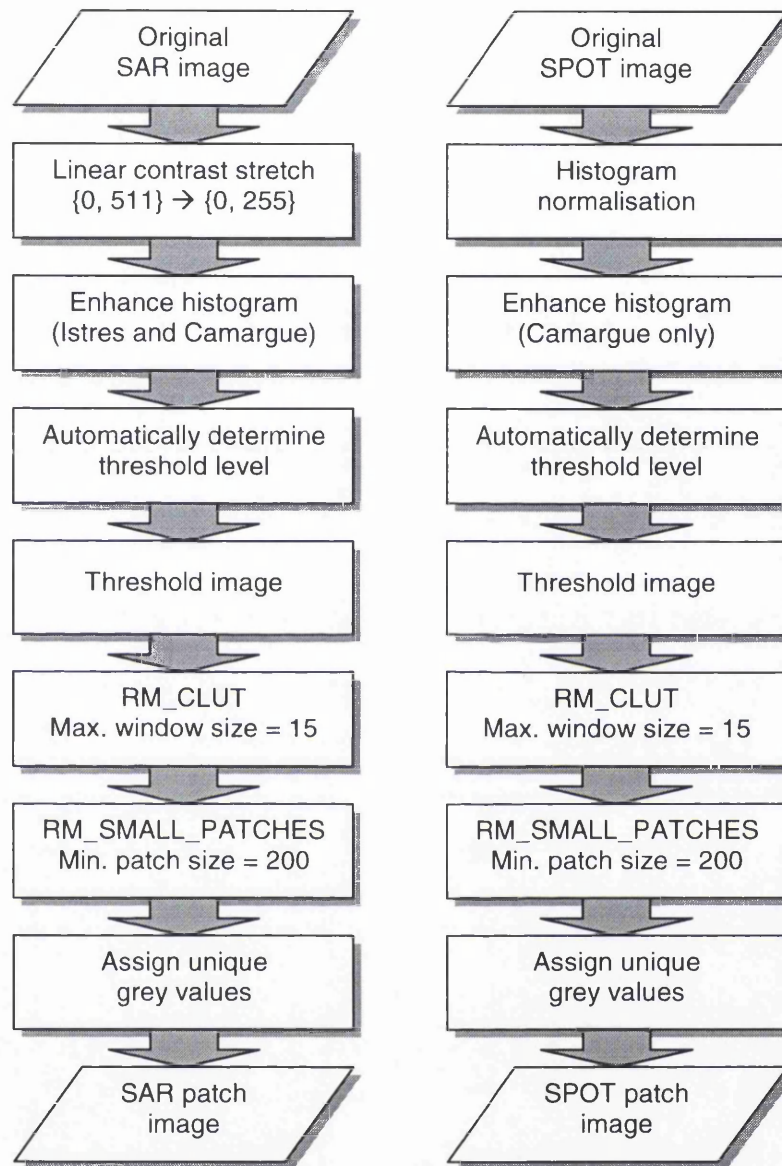


Figure 6.7 Flow chart summarizing automatic thresholding

6.2.3 Homogeneous patch extraction

As with the automatic thresholding, the first step in this extraction procedure was the conversion of the original 16 bit SAR images to 8 bit images using a linear contrast stretch; as before, 0 was assigned to 0 and 511 was assigned to 255. Again the raw SPOT image was normalized to improve the contrast. In order to extract patches using homogeneous patch extraction, a tolerance parameter has to be set by the user. After a few trials, the best values for this parameter appeared to be 13 for the SAR images and 5 for the SPOT images. The results from the homogeneous patch extraction algorithm are shown in figures 6.8 and 6.9 for Istres and figure 6.10 and 6.11 for Camargue.

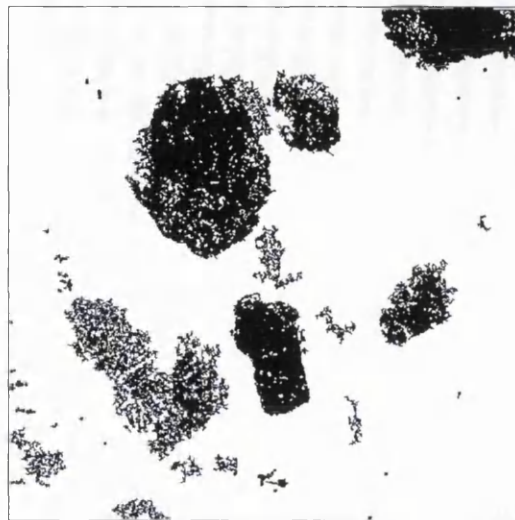
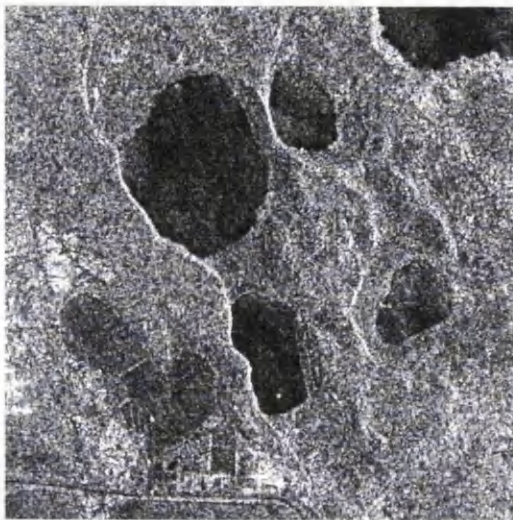


Figure 6.8 Homogeneous patch extraction results for SAR image of Istres

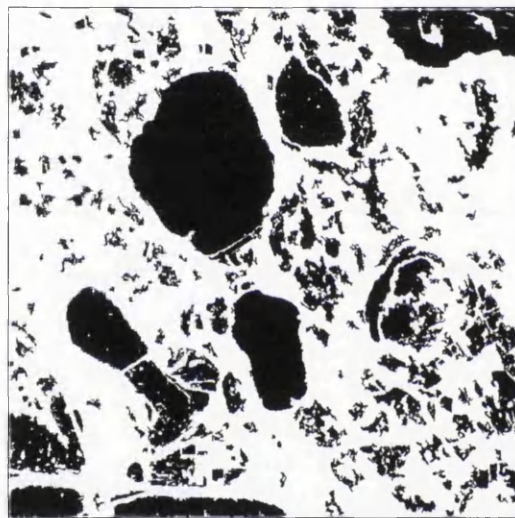
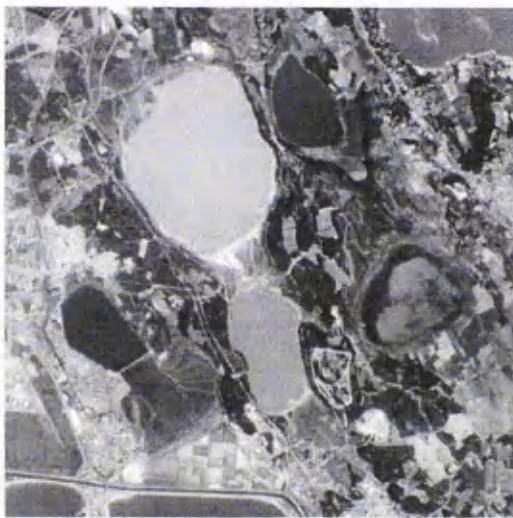


Figure 6.9 Homogeneous patch extraction results for SPOT image of Istres

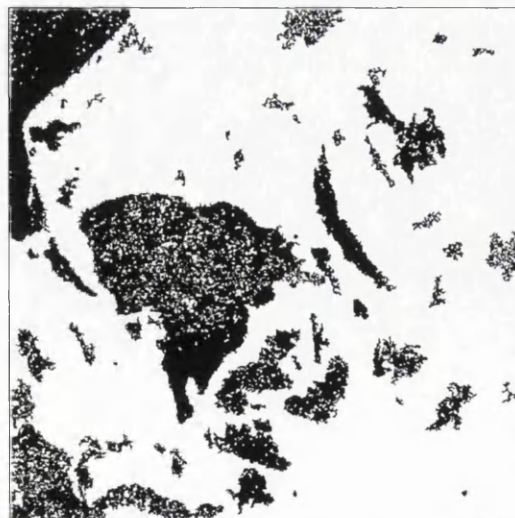
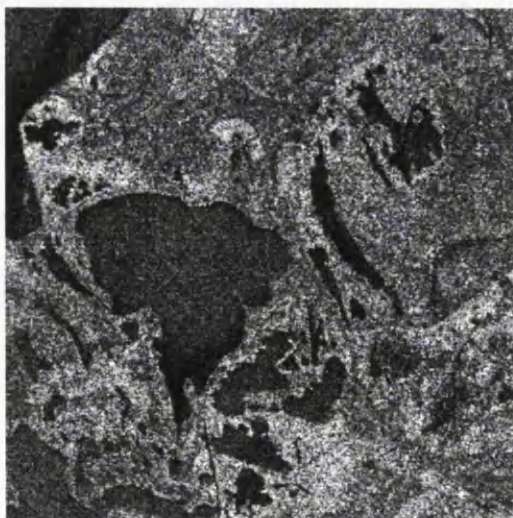


Figure 6.10 Homogeneous patch extraction results for SAR image of Camargue

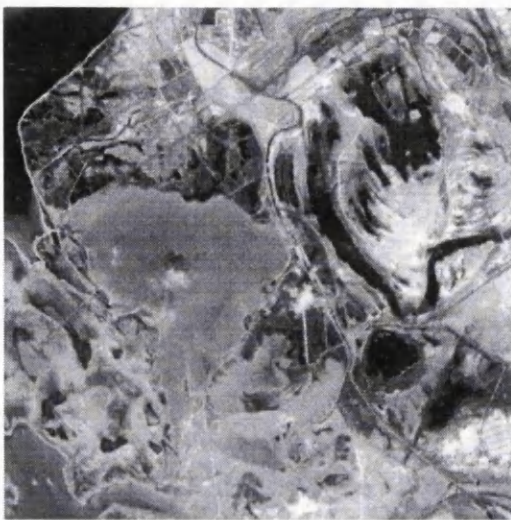


Figure 6.11 Homogeneous patch extraction results for SPOT image of Camargue

Clearly a lot of post-processing is required to fill the holes in the patches extracted from the SAR images, and a lot of clutter has to be removed from the SPOT images. The same two step procedure as used for the automatic thresholding was used here, and the corresponding results are shown in figure 6.12 and figure 6.13.

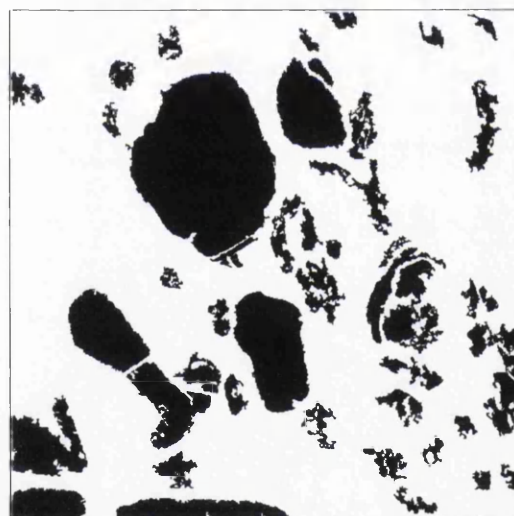


Figure 6.12 Post-processed patches extracted using homogeneous patch extraction for SAR (left) and SPOT (right) images of Istres

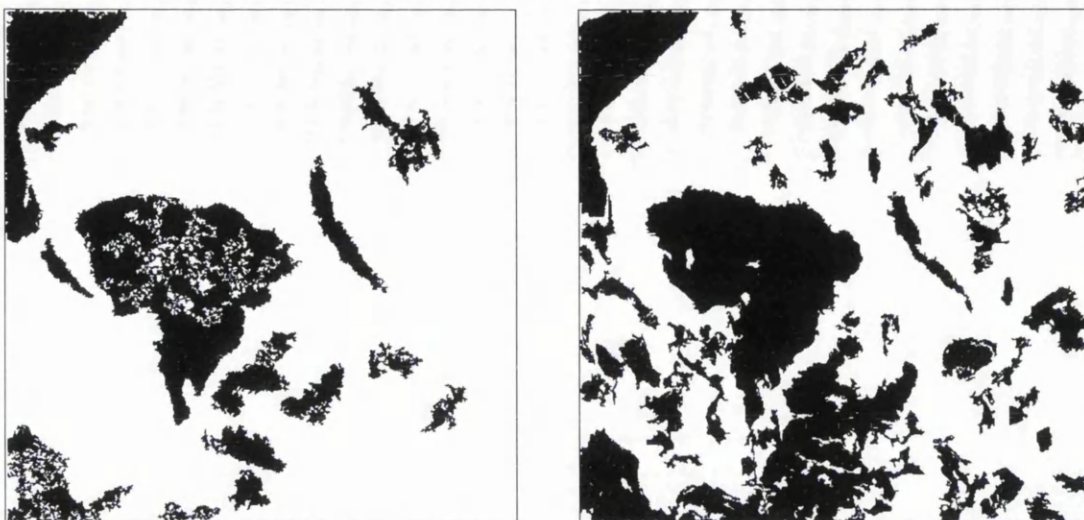


Figure 6.13 Post-processed patches extracted using homogeneous patch extraction for SAR (left) and SPOT (right) images of Camargue

An examination of the above images clearly shows that this method of extracting patches has produced results, which in the case of the SAR images, are quite similar to those produced by automatic thresholding, but in the case of SPOT, are quite different. The homogeneous patch extraction results for the SPOT images of Istres and Camargue are far superior results to those generated using automatic thresholding. The homogeneous patch extraction procedure is summarized in figure 6.14.

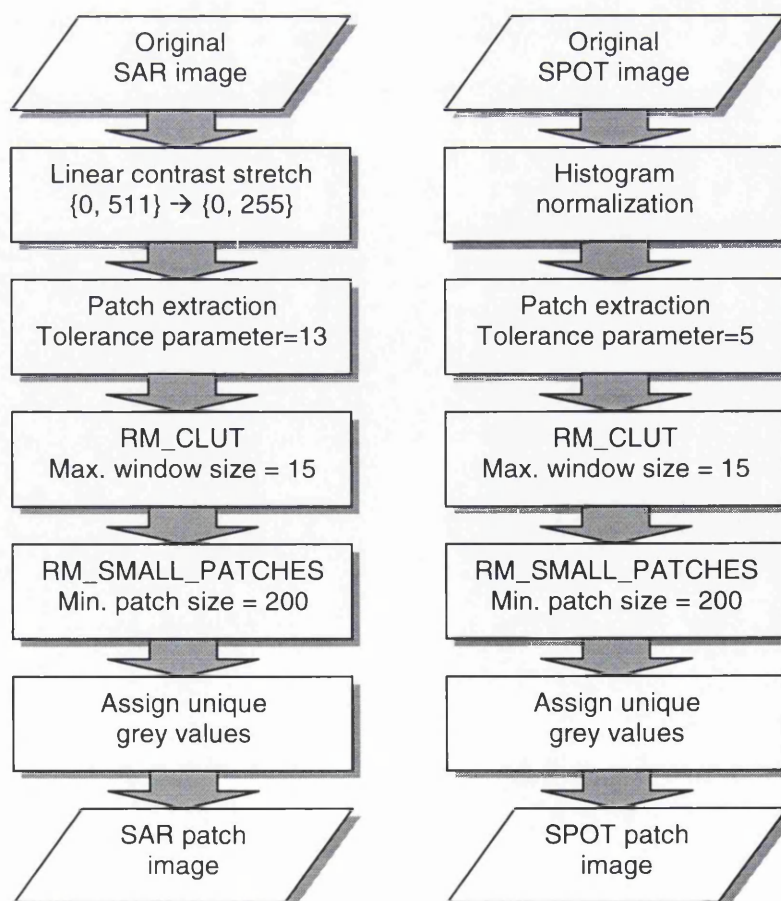


Figure 6.14 Flow chart summarizing homogeneous patch extraction

6.2.4 Segmentation

This section shows the results of applying the segmentation algorithms described in the previous chapter to the SAR and SPOT test images. Since there are four different segmentation algorithms being tested the layout is slightly different to the two previous sections: each algorithm is fully tested before proceeding onto the next.

With the MUM segmentation algorithm there was no need to convert the data from 16 bit to 8 bit since the algorithm works well with both (even though there is a large time saving by using 8 bit data). Experiments were performed with different values of the probability of false alarms parameter p (where the probability of false alarms is given by 10^{-p}) for each of the test images to obtain the best result. The results for the Istres image are shown in figure 6.15, and the results for the Camargue image are shown in figure 6.16.

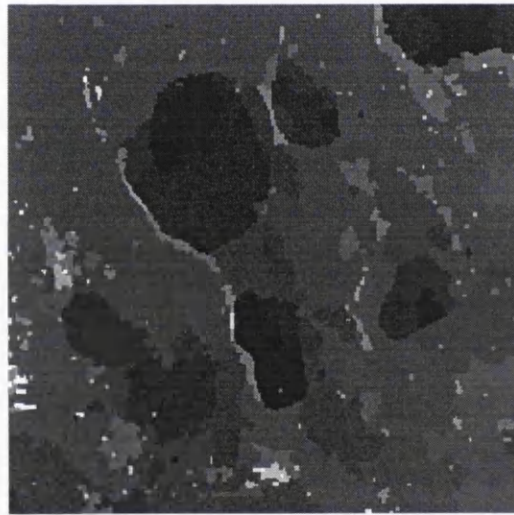
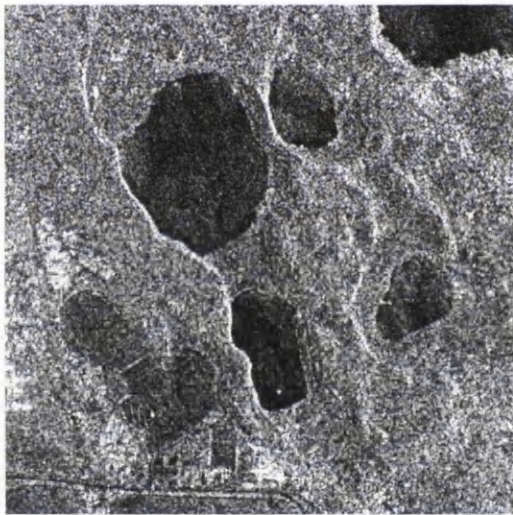


Figure 6.15 Istres SAR image segmented using MUM algorithm with $p = 6$

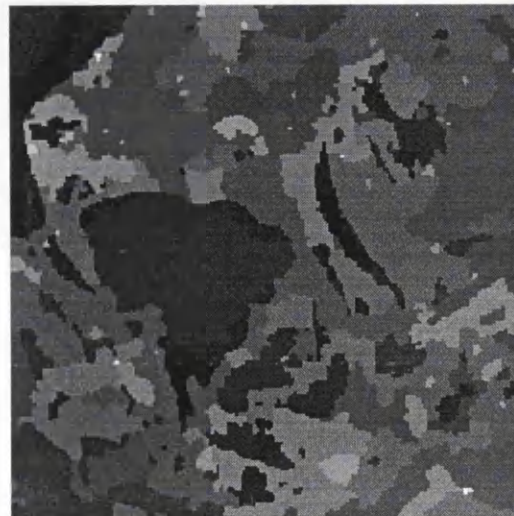
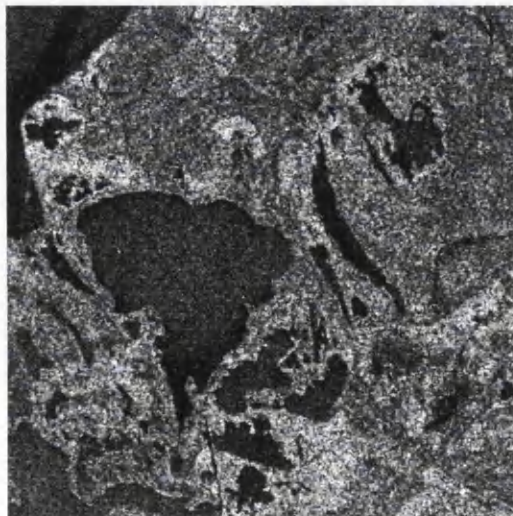


Figure 6.16 Camargue SAR image segmented using MUM algorithm with $p = 6$

The chosen value of p seemed to give the best extraction of features from the background. Although in both images the background has merged in a number of places and the foreground features (the water bodies) remain distinct from the background, a degree of post-segmentation was required to ensure good separation of the features. Patches were merged solely on grey value: neighbouring patches with a grey level difference of less than 30 were merged.

Post-processing of these images is required to remove small insignificant patches, after which holes left behind in the large patches (by the removal of small patches) must be filled. Therefore, for the post-processing of segmented images, the REMOVE_SMALL_PATCHES algorithm must be applied first, and the RM_CLUT algorithm applied second, the reverse of the method used for the previous results. The post-processed images of Istres and Camargue are shown in figure 6.17.

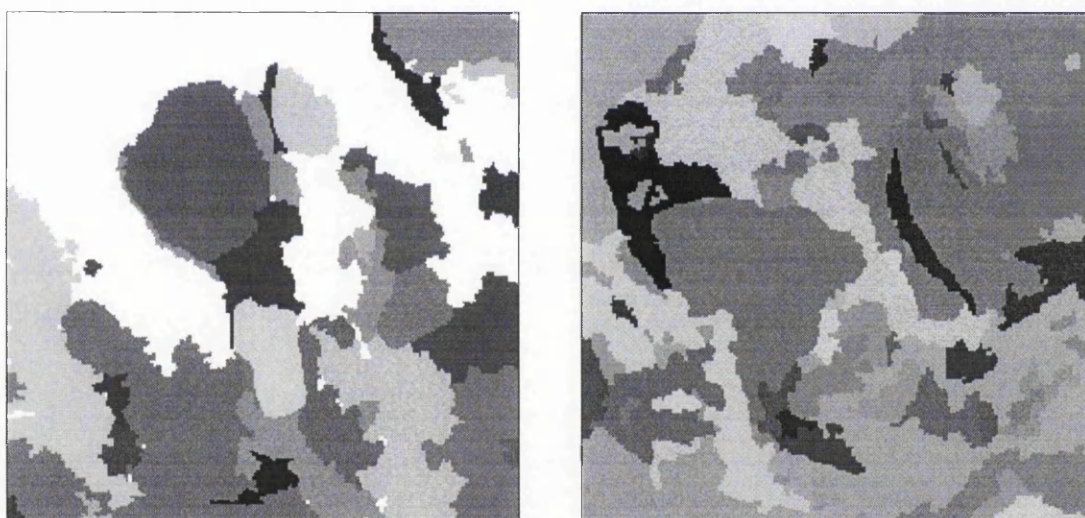


Figure 6.17 Patches extracted from Istres (left) and Camargue (right) using MUM segmentation

In figure 6.17 the patches have been assigned random grey values to clarify the result, but even so it is still a bit difficult to see exactly which features have been extracted. In the Istres image, a least four water bodies, and some other features, have been extracted, and in the Camargue image most of the principal features seem to have been extracted. Many of the patches have been merged by the MUM algorithm. This could be avoided by setting a lower value for the parameter p , but in doing so the patches which represent features would have been split into multiple patches. Therefore, the difficulty with the MUM algorithm is finding a value of p which gives a good balance between splitting of patches and merging of patches.

To achieve the best Istres result (the one shown above) the parameter p was set to 6, the minimum patch size was set to 200 and the maximum window size for the RM_CLUT algorithm was set to 21. For the Camargue image, p was set to 6, the minimum patch size was set to 200 and the maximum window size was set to 21. A flow chart summarizing the MUM segmentation process is shown with the RWSEG process below in figure 6.21.

As with the MUM algorithm, the RWSEG algorithm will also produce results for 16 bit images, so no preprocessing of the image was required. However, the RWSEG algorithm is more complex in that two parameters have to be set (e and j) compared to just the one for the MUM algorithm. This meant that it is much more difficult to find the correct combination of parameters to give the best result. After experimenting with different values, the results shown in figures 6.18 and 6.19 were achieved.

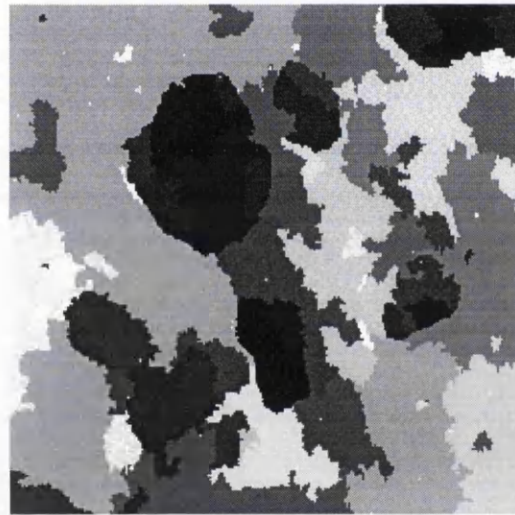
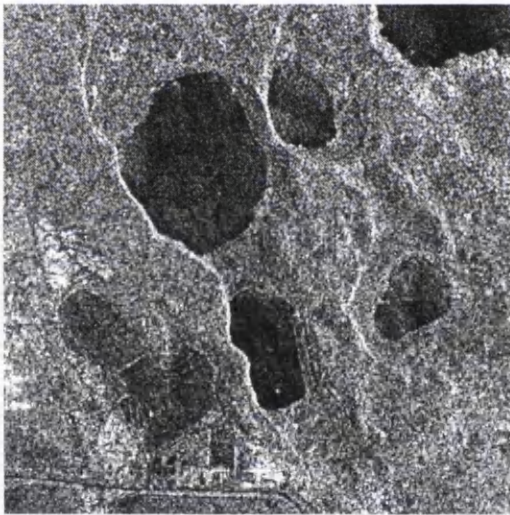


Figure 6.18 Istres SAR image segmented using RWSEG with $e = 3.29$ and $j = 6$

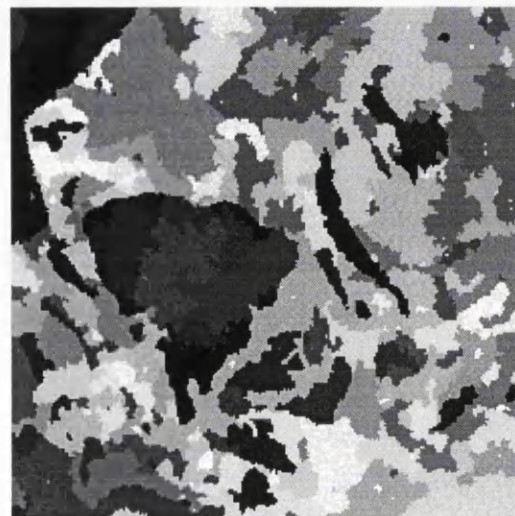
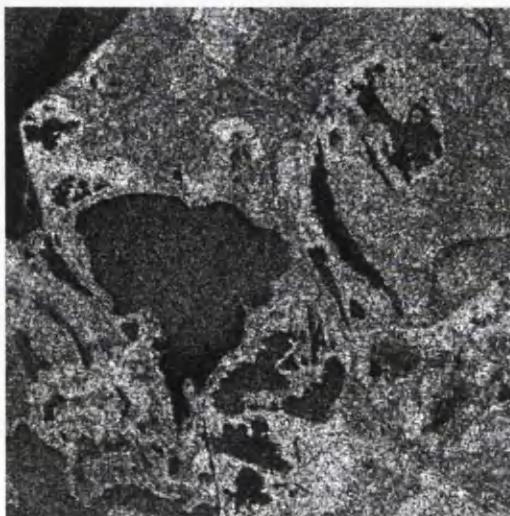


Figure 6.19 Camargue SAR image segmented using RWSEG with $e = 3.29$ and $j = 4$

It proved slightly more difficult to produce good results using the RWSEG algorithm for the two test images than the MUM algorithm. This is because the RWSEG algorithm tends to over-segment the images much more than other algorithms so post-segmentation merging of the patches is required. The algorithm does merge patches to an extent (controlled by the j parameter) but it was found, as with the MUM algorithm, that even more merging was required. This was achieved by merging patches in the resulting image based on grey value alone. For both the Istres and Camargue images, neighbouring patches with a grey level difference of less than 30 were merged.

As with the MUM results, the RWSEG results need to be post-processed. The same method as above was used, i.e. applying the REMOVE_SMALL_PATCHES algorithm first (with minimum patch size set to 200) and the RM_CLUT algorithm second (with the maximum window size set to 21). The results are shown below in figure 6.20. (Note that white regions actually represent areas where no patches exist.)

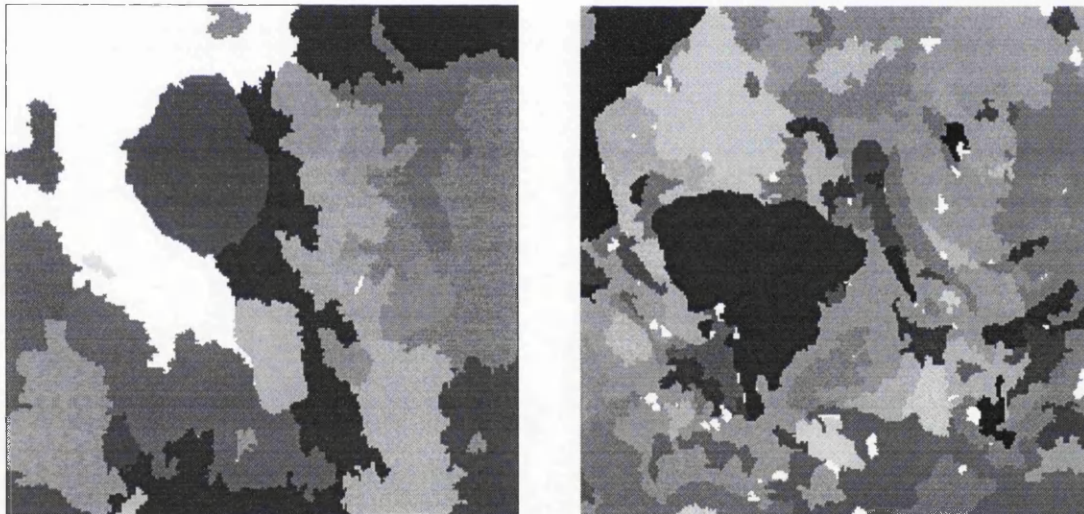


Figure 6.20 Patches extracted from Istres (left) and Camargue (right) using RWSEG segmentation

The final results of the RWSEG algorithm are not that different from the final results of the MUM algorithm. However, they do differ slightly in that the patches are a little more merged in the RWSEG results. Most of the major features seem to have been extracted, but the quality of the result can only be assessed when it is used for matching.

A flow chart summarizing the extraction of patches using RWSEG is shown in figure 6.21 with a flow chart summarizing the MUM algorithm.

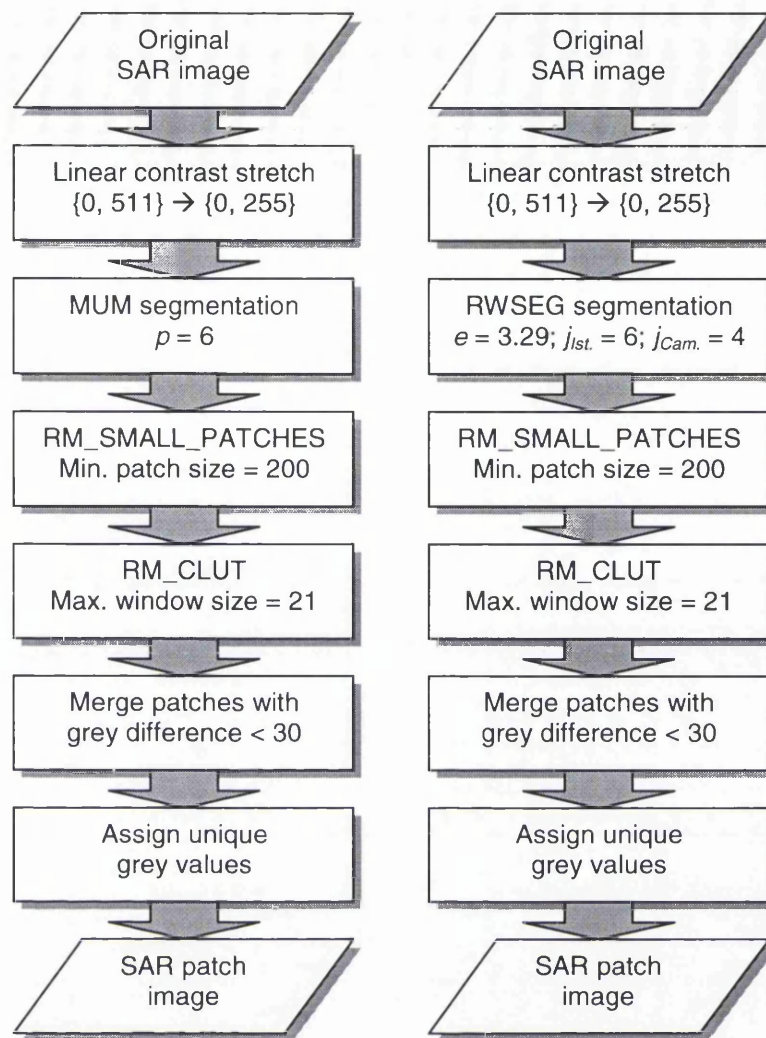


Figure 6.21 Flow charts summarizing MUM and RWSEG algorithms

The next two algorithms, REGSEG and OPTISEG, are used for segmentation of optical images, in particular SPOT imagery. Before either of the algorithms were applied, the SPOT test images were preprocessed. Firstly the images were normalized to improve the contrast of the images, and secondly a smoothing/edge enhancement algorithm (the MHN filter) was applied to the images. For the REGSEG algorithm only one parameter had to be set. As with the other algorithms tested, the values for this parameter for each test image (Istres and Camargue) were found by trial and error, and the results are shown below in figures 6.22 and 6.23.

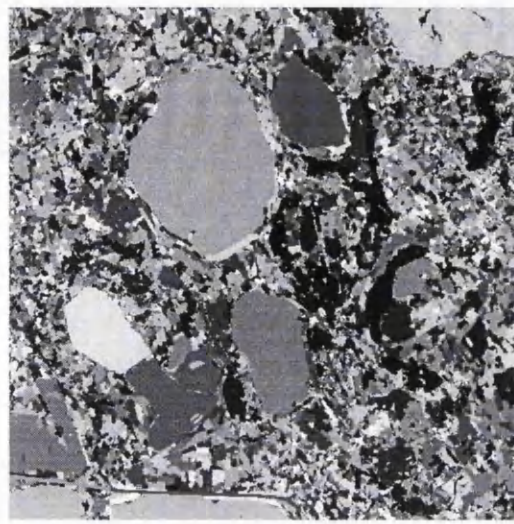


Figure 6.22 Istres SPOT image segmented using REGSEG with $d = 3$

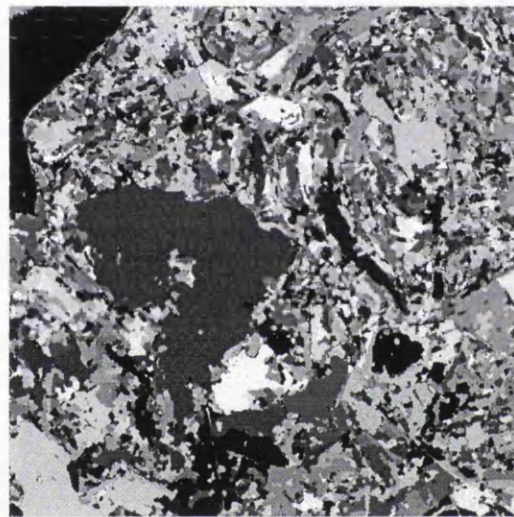
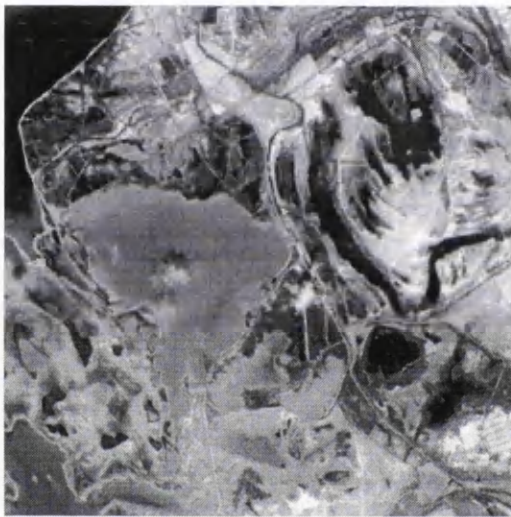


Figure 6.23 Camargue SPOT image segmented using REGSEG with $d = 3$

As mentioned in the previous chapter, the REGSEG algorithm has a propensity to over-segment images. The backgrounds of both the test images have clearly been over-segmented, but the features are quite clear in the foreground. Increasing the value of the parameter d lead to less over-segmentation of the background, but the foreground features became merged with the background. Therefore a value of 3 for d seemed to give the best result in both cases.

Since the backgrounds of the images are clearly over-segmented the small patches need to be removed to leave the large ones, and holes may need to be filled. The same method of clutter removal is used here as was used for the MUM and RWSEG algorithms above. The minimum patch size for REMOVE_SMALL_PATCHES was set as 200 and the maximum window size for RM_CLUT was set as 21. The post-processed results are shown below in figure 6.24.

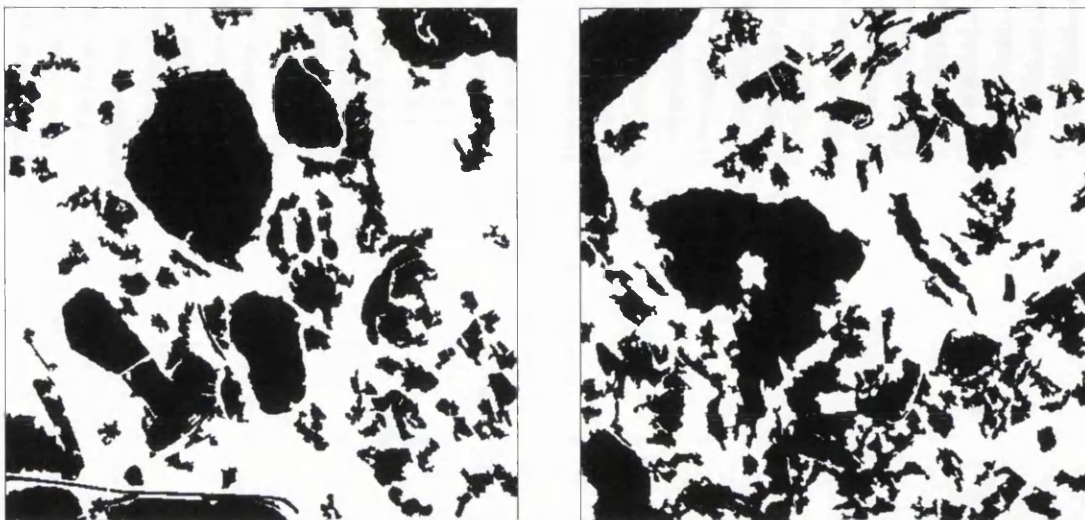


Figure 6.24 Patches extracted from Istres (left) and Camargue (right) using REGSEG segmentation

The above images show that although the initial results of the REGSEG algorithm may have not appeared to be very useful, the algorithm has in fact extracted a large number of useful patches from both images; the post-processing has clearly performed well. A flow chart summarizing the REGSEG segmentation procedure is shown in figure 6.28.

Just as the RWSEG feature extraction procedure was more complicated than the MUM feature extraction procedure, so the OPTISEG feature extraction procedure is more complicated than the REGSEG feature extraction procedure. OPTISEG requires three parameters to be set, so a lot of experimentation was required before good results could be obtained. These results are shown in figures 6.25 and 6.26 below.

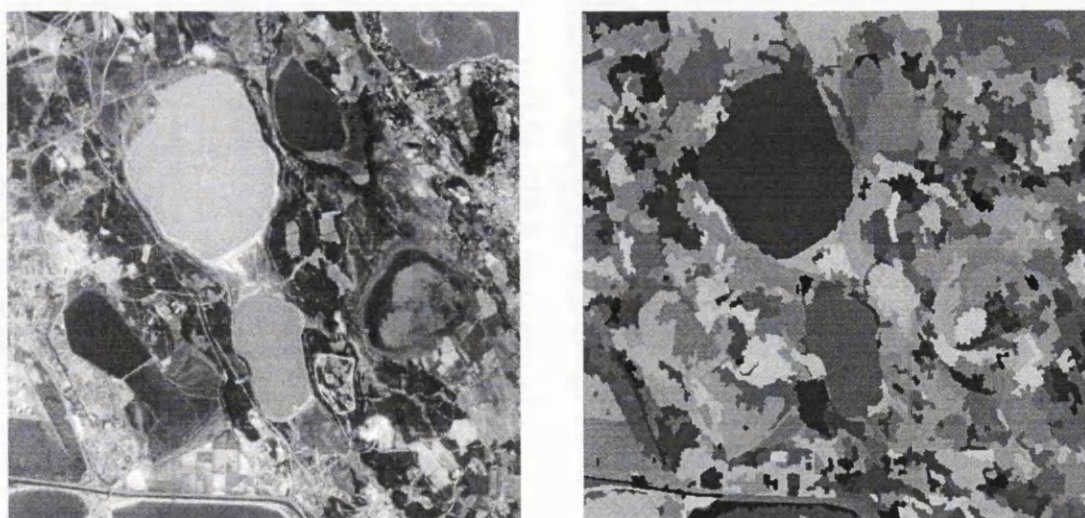


Figure 6.25 Istres SPOT image segmented using OPTISEG with $T = 5$, $\mu = 0.05$, $\nu = -0.2$

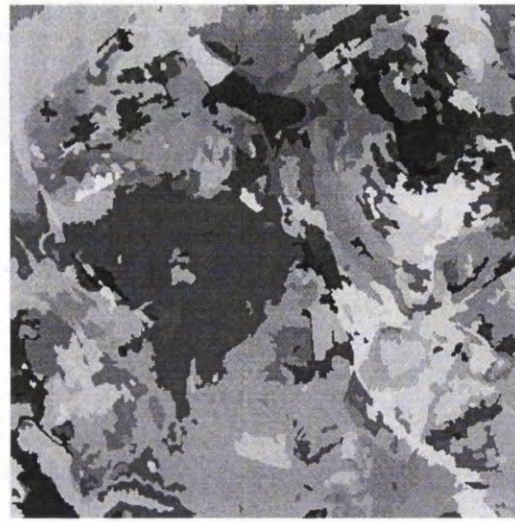
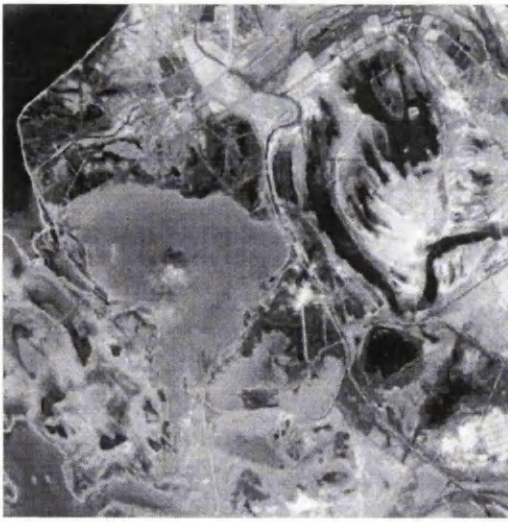


Figure 6.26 Camargue SPOT image segmented using OPTISEG with $T = 5$, $\mu = 0.2$, $\nu = -0.2$

The results obtained using OPTISEG are quite different from those obtained using REGSEG. The REGSEG algorithm gave a result with a greatly over-segmented background, whereas OPTISEG did not over-segment or under-segment either the foreground or background features. The parameters of the OPTISEG algorithm were selected so that the larger features in the could extracted, but this was at the expense of some of the smaller features, which were merged in the process. This is not considered a problem since the smaller features are less useful in the matching procedure than the larger ones .

Once again the same clutter removal procedure was applied, with the minimum patch size set at 200 and the maximum window size set at 21. The results are shown in figure 6.27 below, and a flow chart describing the OPTISEG segmentation procedure is shown in figure 6.28.

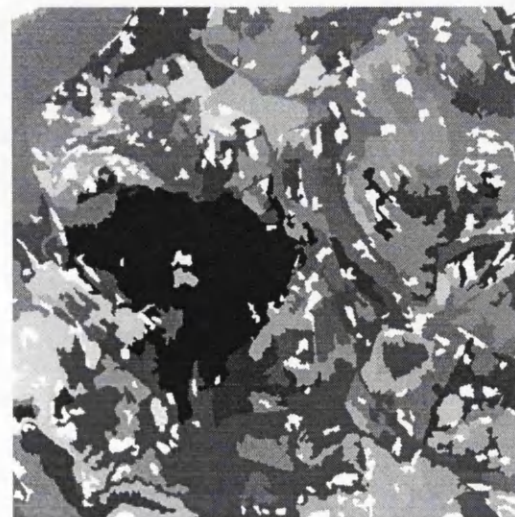
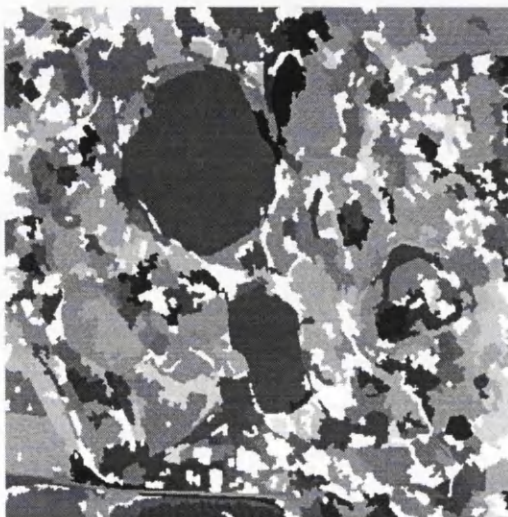


Figure 6.27 Patches extracted from Istres (left) and Camargue (right) using OPTISEG segmentation

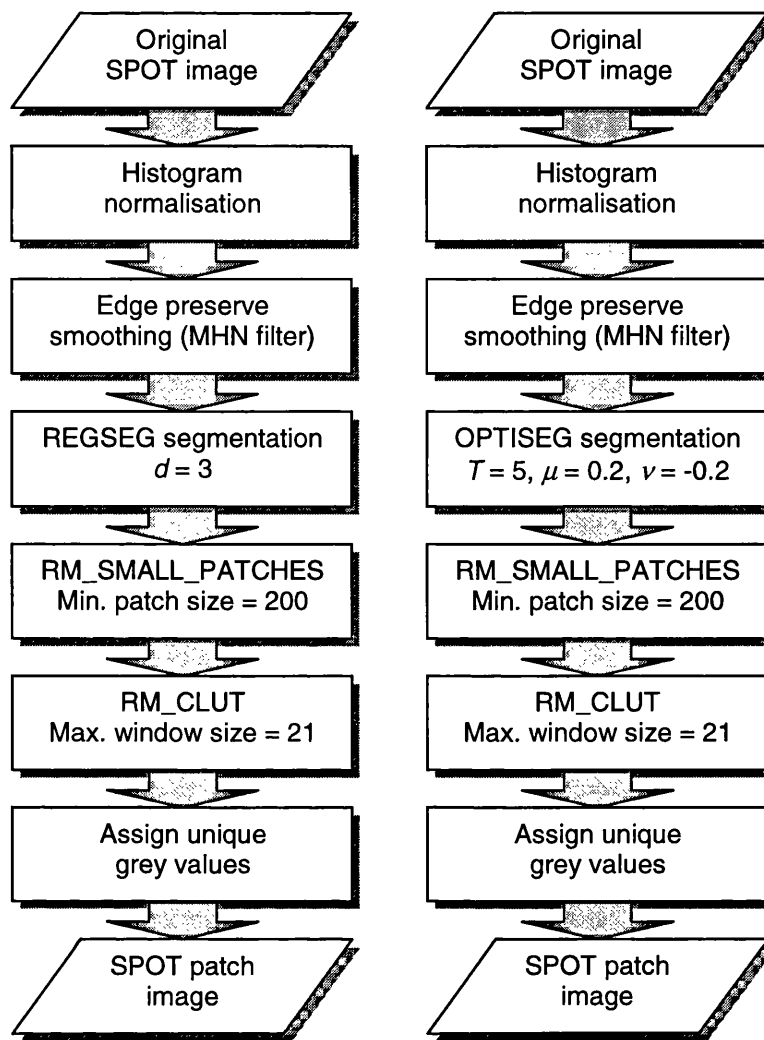


Figure 6.28 Flow chart of optical segmentation procedure

6.2.5 Conclusions

This section has presented the results of applying areal feature extraction algorithms to two SAR and two SPOT test images. It has described in detail exactly how each algorithm has been applied in order to give the best possible results for feature matching. One of the aims of this project is to develop automatic feature extraction algorithms for use in an automatic image registration system, but it is clear from this first section that some of the feature extraction algorithms operate far from automatically. Those algorithms (homogeneous patch extraction and all the segmentation algorithms) have been used with user intervention in order to get the best results. In the context of an automatic system, user intervention would not be possible; this problem is addressed below in § 6.4.2.

The method of automatic thresholding is the only algorithm of those described above which is truly automatic. It requires no user intervention at any stage. However, the above examples show that it does not always produce results useful for patch matching.

Patches which can be used for matching have been extracted from both the SAR and SPOT images of Camargue and the SAR image of Istres, but not from the SPOT image of Istres.

Homogeneous patch extraction has been shown to successfully extract useful patches from all of the images tested above. Since the extraction algorithm only requires one parameter to be set and the post-processing of the results is relatively straightforward, this is a very useful algorithm for the proposed automatic image registration system.

The four segmentation algorithms used in this study all produce results which can be used in the patch matching procedure. The most similar results come from the two SAR segmentation algorithms (MUM and RWSEG). Both give useful patches, but the MUM algorithm only needs one parameter to be set while the RWSEG algorithm needs two to be set. Therefore, although the automation of the setting of the single MUM parameter will be more straightforward than that of the two RWSEG parameters, the latter does offer more flexibility in the type of result which can be achieved. The same is true for the two optical segmentation algorithms (OPTISEG and REGSEG): they both produce useful results, but OPTISEG requires three parameters to be set, compared to just one for REGSEG. Therefore, the automation of the OPTISEG algorithm will be more difficult than the automation of the REGSEG algorithm.

The results shown above illustrate very clearly the need for multiple feature extraction algorithms. A single algorithm would be insufficient to produce enough useful patches for automatic registration.

A final point should be made about the test images used. The SAR images were acquired four years after the SPOT images were acquired, and since some of the major features in the images are water bodies, it would perhaps be expected that there would be some change between the images. However, this does not appear to be the case and the images are quite similar, apart from the cloud contamination that is clearly visible in the SPOT image of Camargue.

6.3 Patch matching

6.3.1 Introduction

So far this chapter has presented the results of extracting patches from four remotely sensed test images. The next stage of the image registration procedure is to code and match these patches. Once matched they can be used to register the corresponding

images. This section describes the patch matching procedure used in this system of automatic image registration. Although the method presented here has a few similarities to that described by Morgado and Dowman (1997) and Abbasi-Dezfouli and Freeman (1994), it is intrinsically new.

The first section below describes how the patches are coded. A number of attributes are determined for each patch, such as area, perimeter length and shape. The second section below describes how the patches are matched based on their attributes, and how that matching is refined. The final section uses the matching results to perform the registration, and the results are presented.

6.3.2 Patch attribute encoding

This section describes the methods used to encode the attributes of the patches which are used in the matching procedure. These attributes are:

- ❑ area;
- ❑ perimeter length;
- ❑ bounding rectangle; and
- ❑ location in the image.

The area of each patch is simple to calculate since it merely involves counting up all the pixels which make up that patch (see figure 6.29 below). This is trivial if all the patches in the image have a different grey value. The post-processing of the extracted patches featured an algorithm for unique grey value assignment, so determining the area of each patch is therefore straightforward. The perimeter length of each patch is determined by counting up all the pixels in the perimeter of the patch, and the bounding rectangle is calculated by determining the extent (in rows and columns) of the patch (see figure 6.29 below).

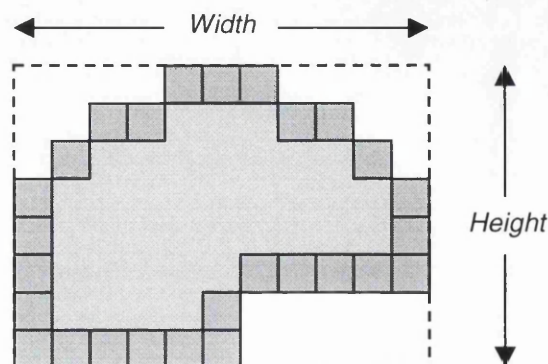


Figure 6.29 Area, perimeter and bounding rectangle of typical patch

In the above diagram, the grey pixels (both light and dark) represent the whole area of the patch. The dark grey pixels are the pixels which make up the perimeter. By simply counting the dark grey pixels it is possible to determine the length of the perimeter. The bounding rectangle is expressed in terms of width and height (shown in the above diagram) and measured in pixels.

The last attribute of the patch to be determined is the location of the patch in the image. This is calculated by measuring the co-ordinates of the centroid of the patch. The centroid is calculated by simply taking the mean of the row co-ordinates of all the pixels in the patches, and the mean of the column co-ordinates of all the pixels. Although comparison of centroid locations is not a reliable way of matching patches, it is used in the matching procedure to refine the matching results.

One further attribute that previous studies have considered useful in patch matching, but which has not been used here, is the chain code of the boundary of the patch. Chain coding (Freeman, 1961) is a technique used to describe the changing direction of the boundary, and therefore describes the shape of the patch. In this study where patches extracted from SAR and optical images are being matched, chain codes would not necessarily increase the accuracy of the matching since, although the general shapes of corresponding patches are similar, their boundaries can be quite different. For an example of this, compare the corresponding patches in figure 6.12 above. Furthermore, there are other problems associated with chain codes which can make it difficult to match patches successfully. These problems are discussed in Abbasi-Dezfouli and Freeman (1994).

This discussion of attribute coding of patches has only featured five different types of attributes. However, there are numerous other methods of describing the shapes of patches, such as polygonal approximations, signatures and boundary segments. These are not considered here since the method of matching described below is simple, accurate and robust, and so their inclusion would be superfluous. A detailed discussion of these other methods of patch coding can be found in Gonzalez and Woods (1992).

The attributes described above are exemplified below using the patches extracted from the SPOT image of Istres using homogeneous patch extraction (shown in figure 6.30). There are 31 patches in this image – the attributes of the four largest patches are listed in table 6.1 below. The reference numbers in the image refer to the corresponding reference numbers in the table.

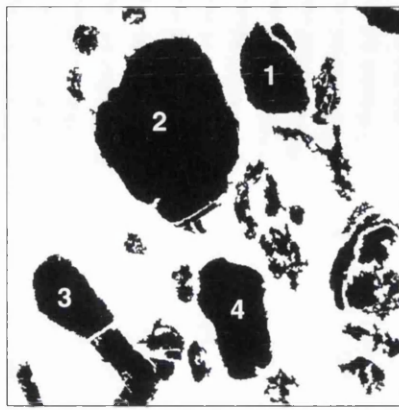


Figure 6.30 Patches from SPOT image of Istres with corresponding reference numbers

Reference number	Area	Perimeter length	Centroid (column)	Centroid (row)	Bounding rectangle rows	Bounding rectangle columns
1	4291	260	266.917	68.799	92	69
2	18978	549	164.209	122.386	186	149
3	4097	255	63.095	291.865	84	81
4	6356	330	231.220	309.265	122	75

Table 6.1 Attributes of patches in SPOT image of Istres processed with homogeneous patch extraction (all units are pixels)

The coding of the patches is a completely automatic process requiring no human intervention. The patches extracted from the images (described in the previous section) were all coded in this manner.

6.3.3 Matching

This section describes the process used to match the patches extracted from the SAR and SPOT images. Although it is similar to previous work, the procedure is essentially new. It works in a stepwise manner where the matching results are refined in each step. The first step of the matching algorithm matches all the patches from the first image (image 1) to patches in the second image (image 2), and *vice versa*. This is done by using the differences in the attributes of the patches to determine a cost function which is minimized to give the best match. The cost function is also used to removed multiple matches. The second step refines the results by eliminating false matches by comparing the shapes of matched patches. The third step further refines the matching results by analysing the separations of matched patches. This matching procedure is summarized in figure 6.31, and the three steps are described in detail in the remainder of this section.

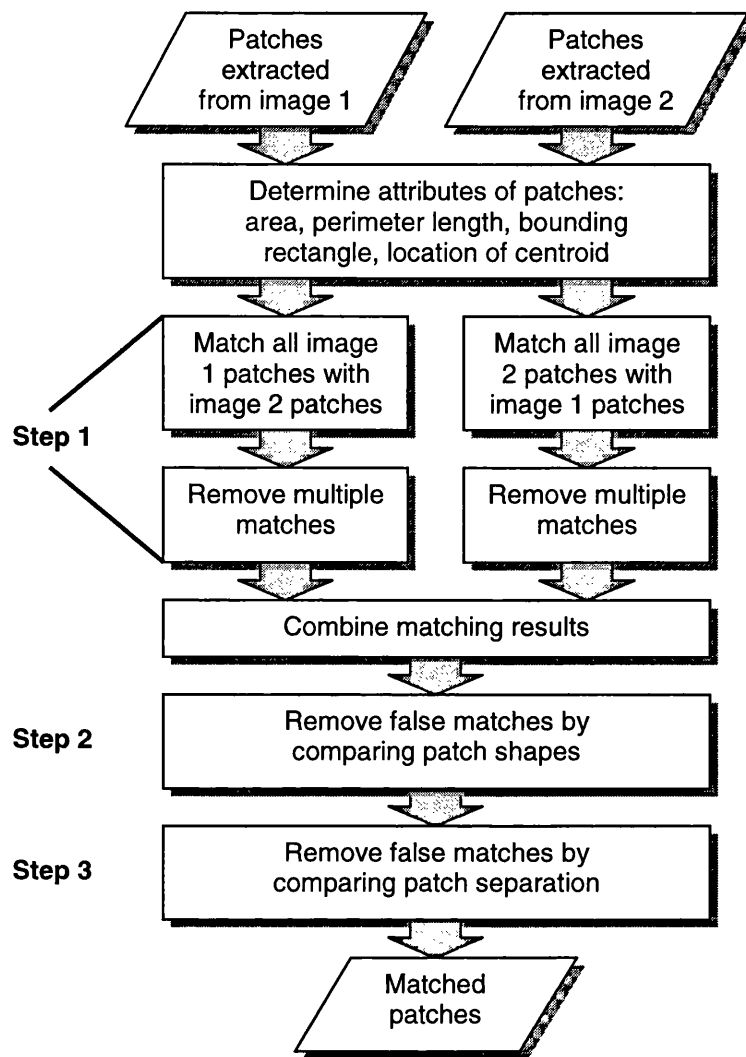


Figure 6.31 Flow chart summarizing patch matching procedure

Step 1 – Cost function matching

The use of a cost function for image matching was described by both Morgado and Dowman (1997) and Abbasi-Dezfouli and Freeman (1994). However both of these authors found that there are some problems associated with its implementation. A general point was that on its own it will not differentiate all the good matches from the bad matches. This is because incorrect matches sometimes give a low value for the cost function, and correct matches sometimes give high values. Therefore, more information has to be used to improve the matching results, such as using a chain code method to describe the shapes of the patches, or using the relative geometry of the matched patches with respect to each other in order to highlight poor matches. Although it was shown that both of these methods worked in the examples cited by the authors, they would not necessarily work here. The reason for not using chain codes method was given in § 6.3.2. The problem with using the relative geometry of matched patches is that there may not be enough good matches to ensure a reliable result could be determined, or too many bad but geometrically similar matches could easily lead to an

incorrect result. Therefore, there is a question of robustness associated with these two methods of improving the match results.

The cost function used by Morgado and Dowman (1997) is shown below in equation 6.1.

$$\Gamma = \frac{1}{2} |a_1 - a_2| + |p_1 - p_2| + |r_1 - r_2| + |c_1 - c_2| + \dots \quad (6.1)$$

where Γ is the value of the cost function, a_i is the area of patch i , p_i is the perimeter length, and r_i and c_i are the height (in row pixels) and width (in column pixels) of the bounding rectangle. The last term which has not been shown here represents differences in chain codes.

A specific problem associated with this function is that it was found that the area component influenced the results more than the other attributes. To get around this problem, the value of the area component was halved to reduce its influence. However, since areas and lengths are being compared, it seems more reasonable to take the square root of the area component, so that all the components being compared have the same dimensionality. Furthermore, with the above function, larger patches will always give a larger cost function value than smaller patches. Therefore, to ensure this is not the case, the differences in components have been normalized. Thus, the cost function used in this study is expressed by the equation:

$$\Gamma = \left| \frac{a_1 - a_2}{a_1 + a_2} \right|^{\frac{1}{2}} + \left| \frac{p_1 - p_2}{p_1 + p_2} \right| + \left| \frac{r_1 - r_2}{r_1 + r_2} \right| + \left| \frac{c_1 - c_2}{c_1 + c_2} \right| \quad (6.2)$$

The cost function is determined for the first patch in image 1 and all the patches in image 2. The combination of patches which gives the minimum value is accepted as the best match. The process is repeated for the second, third and all subsequent patches in image 1 until they have all been matched with patches in image 2. In the situation where a patch from image 2 has been matched with two different patches from image 1, the match with the lowest cost function is accepted as the correct one. The result is that all the patches in image 1 have been matched, but not necessarily all of the patches in image 2 have been matched. In order to ensure that all of the patches in image 2 have the opportunity to be matched with all of the patches in image 1, the process is repeated with the order of the images reversed. The first patch in image 2 is matched with all the

patches in image 1, as is the second, third, and so on. Multiple matches are again eliminated using the value of the cost function. Before the matching results can be refined in the second step of the matching procedure, the results from the two processes described above have to be combined.

An example of the cost function matching results, generated using patches extracted from the SAR image of Istres using MUM segmentation and patches from the corresponding SPOT image using REGSEG segmentation, are shown in table 6.2.

SAR patch grey level	SPOT patch grey level	Cost function	Patch separation (pixels)
2	5204	0.463	112.429
3	220	0.484	0.353
7	13897	0.270	225.892
9	3417	0.550	253.794
10	2901	0.277	4.651
12	3754	0.229	6.835
13	18389	0.265	267.599
18	26749	0.240	514.928
21	13746	0.662	209.264
22	21272	0.394	275.525
24	8585	0.259	378.102
28	18558	0.306	166.298
30	21476	0.327	198.795
32	11307	0.313	59.894
33	25114	0.129	242.566
37	21579	0.137	303.653
39	20029	0.286	395.225
43	16449	0.549	43.834
45	15793	0.287	3.995
56	1339	0.414	330.546
60	11922	0.416	382.007
61	16119	0.474	123.264
71	17966	0.215	347.164
73	13154	0.714	222.182
75	6637	0.276	384.673
76	16520	0.186	386.544
77	4572	0.382	418.060

Table 6.2 Example of cost function matching results

Table 6.2 shows that there are a large number of matches, but the majority of these are false (this is clear from the patch separation distances). The next step in the matching procedure refines these results and eliminates false matches by comparing the shapes of the matched patches.

Step 2 – Refinement by comparing patch shape

Previous studies have used chain coding to match patches based on shape. However, this procedure will not necessarily produce reliable results for patches extracted from SAR and SPOT imagery. Therefore, a much simpler method is introduced here. To detect a false match, a pair of matched patches are laid on top of each other with their centroids coinciding, and the number of pixels that they have in common are counted. Patches with similar shapes will have a large number of pixels in common, whereas patches with different shapes will only have a few pixels in common. In order that results from different pairs of patches can be compared, the results are expressed as a ratio of pixels in common to total number of pixels in the two patches. By selecting a threshold for minimum amount of overlap it is simple to separate the correct matches from the false matches. It was decided to use a threshold of 70% for all the results. Table 6.3 shows the matching results from table 6.2 with correct matches highlighted according to their percentage of overlap.

Obviously it is possible that some false matches may still be accepted as true matches, but now the correct matches outnumber the false matches. The third stage of the matching procedure further refines the results to ensure that no false matches remain.

SAR patch grey level	SPOT patch grey level	Cost function	Patch separation (pixels)	Percentage overlap
2	5204	0.463	112.429	19.4
3	220	0.484	0.353	88.5
7	13897	0.270	225.892	33.1
9	3417	0.550	253.794	31.2
10	2901	0.277	4.651	89.9
12	3754	0.229	6.835	94.4
13	18389	0.265	267.599	19.9
18	26749	0.240	514.928	37.2
21	13746	0.662	209.264	66.8
22	21272	0.394	275.525	36.6
24	8585	0.259	378.102	40.1
28	18558	0.306	166.298	49.3
30	21476	0.327	198.795	52.6
32	11307	0.313	59.894	56.1
33	25114	0.129	242.566	56.1
37	21579	0.137	303.653	65.2
39	20029	0.286	395.225	57.3
43	16449	0.549	43.834	30.7
45	15793	0.287	3.995	88.7
56	1339	0.414	330.546	23.0
60	11922	0.416	382.007	62.9
61	16119	0.474	123.264	59.6
71	17966	0.215	347.164	51.4
73	13154	0.714	222.182	29.0
75	6637	0.276	384.673	48.2
76	16520	0.186	386.544	35.1
77	4572	0.382	418.060	47.1

Table 6.3 Refined matching results with correct matches highlighted

Step 3 – Refinement using patch separation

The third and final stage of the matching process removes the remaining false patches (if there are any) by comparing the patch separation distances for all the matched patches. Since the images being matched are initially aligned with each other, it true to say that the patch separation for all correctly matched patches should be very similar. By plotting a histogram of all the patch separations, a cluster will exist which represents the correct matches. The patch separation for the false matches will distributed randomly. Therefore, by eliminating all the matches with a patch separation distance outside that cluster, only the correct matches remain. For this method to be reliable, there must be a significant number of correctly matched patches in order to identify the cluster.

The final results with false matches eliminated are shown in table 6.4, and the matched patches are shown in figure 6.32. The images confirm that all of the matches are correct.

SAR patch grey level	SPOT patch grey level	Cost function	Patch separation (pixels)	Percentage overlap ($\times 100$)
3	220	0.484	0.353	0.885
10	2901	0.277	4.651	0.899
12	3754	0.229	6.835	0.944
45	15793	0.287	3.995	0.887

Table 6.4 Final matching results

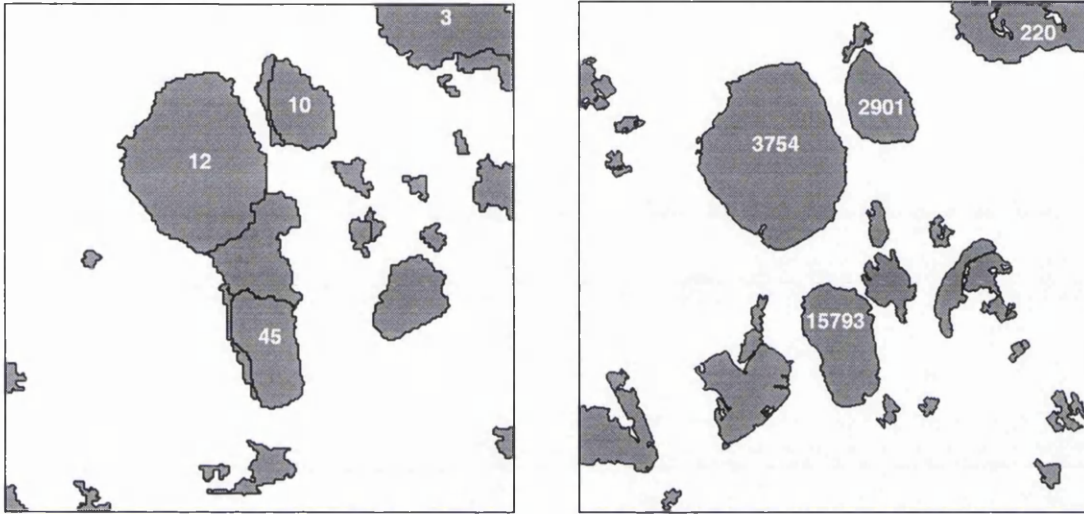


Figure 6.32 Matched patches from SAR and SPOT images of Istres

The above results clearly illustrate that the method of patch matching proposed here has been successful. Table 6.4 shows that comparison of the cost function alone for different combinations of patches is not a suitable method of patch matching. However, when it is used in conjunction with shape information, in the way described above, accurate results can be obtained. Furthermore, it should be pointed out that the proposed method of patch matching is simple, quick, efficient and automatic. No user intervention is required at any stage.

This matching procedure was applied to all the combinations of patches extracted from SAR and SPOT images. The results are shown in table 6.5.

Method of patch extraction	Camargue		Istres	
	Total matches	Total correct matches	Total matches	Total correct matches
SAR auto thresholding SPOT auto thresholding	46	1	50	0
SAR auto thresholding SPOT homogeneous patch	82	0	59	3
SAR auto thresholding SPOT REGSEG	150	2	108	3
SAR auto thresholding SPOT OPTISEG	206	1	333	3
SAR homogeneous patch SPOT auto thresholding	39	1	50	0
SAR homogeneous patch SPOT homogeneous patch	75	0	59	3
SAR homogeneous patch SPOT REGSEG	143	2	108	2
SAR homogeneous patch SPOT OPTISEG	199	0	333	3
SAR MUM SPOT auto thresholding	94	2	83	0
SAR MUM SPOT homogeneous patch	130	1	92	4
SAR MUM SPOT REGSEG	198	4	141	4
SAR MUM SPOT OPTISEG	254	0	366	2
SAR RWSEG SPOT auto thresholding	126	1	73	0
SAR RWSEG SPOT homogeneous patch	162	0	82	5
SAR RWSEG SPOT REGSEG	230	2	131	3
SAR RWSEG SPOT OPTISEG	286	1	356	3

Table 6.5 Summary of all matching results for Istres and Camargue

The results in table 6.5 are quite interesting. The most striking trend is that the number of possible matches and the number good matches are very dependent upon which image and which combination of feature extraction algorithms are used. The fact that it is not possible to select one combination of algorithms which will perform well on all

pairs of images justifies the use of multiple feature extraction algorithms in this study. Previous studies have overlooked or ignored this fact by selecting images which give good results with the algorithms being used. A couple of other conclusions that can be drawn from table 6.5 are firstly that there are more matches for the Istres image than the Camargue image (surprising perhaps when the images are compared) and secondly that the number of possible matches is not a guide for number of good matches. Therefore the combinations of algorithms should be judged only on the number of good matches found.

In conclusion, this section has described a robust and reliable procedure for matching patches extracted from SAR and SPOT imagery, and presented the results of applying this procedure to 16 pairs of feature extraction images. The next section uses these results to the approximately register the corresponding pairs of images.

6.3.4 Image registration

The previous section has shown that the patch matching algorithm described here can successfully match patches extracted from SAR and SPOT imagery. This section demonstrates how those results can be used to register the two pairs of images (Istres and Camargue). Table 6.5 above shows the number of good matches which were found for each combination of feature extraction algorithm. The combinations producing the greatest number of matches will be used to perform the registration. It is not necessarily true that these combinations will give the most accurate registration, but for the purposes for this demonstration, they will give the largest number of residuals.

For the Camargue test site, the greatest number of matches were found using the patches extracted from the SPOT image using the REGSEG algorithm, and the patches extracted from the SAR image using the MUM algorithm. The corresponding matched patches are shown in figure 6.33 below.



Figure 6.33 Matched patches extracted from SPOT (left) and SAR (right) images of Camargue

The centroids of these patches have been marked with small crosses. It is these centroids which are used as tie points to register the images. All four points were used to generate the parameters of an affine transformation. This transformation was used to register the SPOT image to the SAR image. The residuals of the transformed tie points, along with the original and transformed co-ordinates of the centroids are given in table 6.6 below.

The corresponding data is also shown below for the Istres test site. In this case, the combination of feature extraction algorithms which gave the greatest number of tie points was RWSEG for the SAR image and homogeneous patch extraction for the SPOT image. The matched patches are shown in figure 6.34.

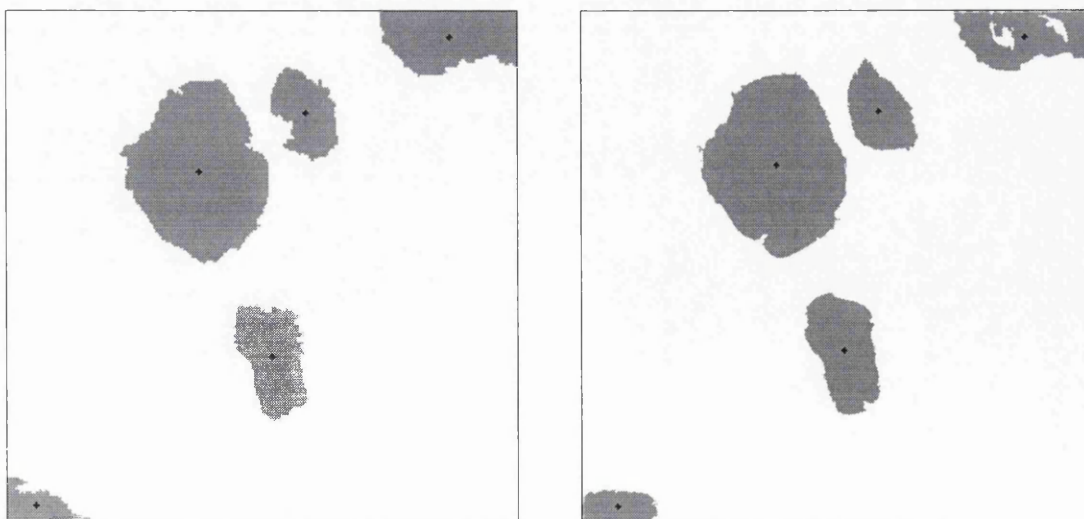


Figure 6.34 Matched patches extracted from SPOT (left) and SAR (right) images of Istres

In this case five patches have been matched, so there are five tie points with which to determine the parameters of the transformation for registering the images. Once again

an affine transformation was used to register the SPOT image to the SAR image. The original co-ordinate data of the tie points and the residuals are given in table 6.6.

	SPOT original		SAR original		SPOT transformed		Residual	
	Column	Row	Column	Row	Column	Row	Column	Row
Camargue	67.73	37.80	78.86	36.82	70.20	37.64	2.47	-0.16
	123.88	399.41	125.20	398.25	122.47	399.50	-1.41	0.09
	266.81	175.52	267.23	176.57	262.62	175.79	-4.19	0.27
	347.89	385.76	351.59	386.69	351.03	385.57	3.13	-0.20
Istres	26.01	444.27	26.92	444.03	23.56	441.94	-2.45	-2.33
	100.80	298.92	102.44	299.10	99.24	300.68	-1.55	1.77
	154.39	196.21	161.14	192.34	158.08	196.60	3.69	0.39
	341.26	263.22	347.64	265.43	344.80	265.04	3.54	1.82
	497.12	36.30	496.42	29.39	493.89	34.67	-3.22	-1.64

Table 6.6 Co-ordinates of Camargue and Istres tie points with residuals

It should be remembered that the residuals in the above table are not an independent check of the accuracy of the registration, but even so they do give an indication of the quality of the registration. However, with so few tie points it is very difficult to interpret very much from these residuals. One important point that can be clearly seen is that there are no blunders in the matching: all of the residuals have the same order of magnitude of less than 5 pixels. A blunder would be apparent since it would lead to some quite large residuals.

6.3.5 Conclusions

This section has presented a method for matching patches extracted from SAR and optical imagery, and shown the types of results that can be achieved. The idea of extracting patches from different types of imagery and matching them in order to register the images is not a new one, but the specific method of doing this described here is new. Although some of the original ideas for matching were based on previous work, it was found that new techniques had to be developed if the patches were to be matched accurately and efficiently. These new techniques have been designed to produce accurate and reliable results automatically, without the need for human intervention, which is essential for a robust automatic image registration system.

To summarize, the patch matching method works by matching all the patches from one image to patches in the corresponding image, using a simple cost function derived from attributes of the patches. Multiple matches are removed, and the results are refined by comparing the shapes and the separations of the matched patches. The result is that for each combination of feature extraction algorithms a number of reliable matches can be

found. The quantity of matches depends on which feature extraction algorithms and which images were used. For the two test images used in this part of the study, it was found that the segmentation algorithms tended to give the most matches, and the thresholding algorithm the least number of matches. However, it is not expected that a similar distribution of matches would be found for other images.

Lastly, this section has shown how the matched patches can be used to derive the parameters of a transformation function which in turn is used to register the images. Although there are too few matches to allow a full analysis of the accuracy of the registration, the residuals do indicate that there were no matching blunders.

Previous researchers which have developed similar techniques of feature based image registration have usually stopped at this point and presented their methods as prototype systems for automatically registering imagery. However, there are still a number of problems which need to be overcome if the method proposed here is to be useful as an automatic image registration system. Some of the issues which need to be addressed are:

- ❑ improving the quantity of matches;
- ❑ improving the spatial distribution of the matches;
- ❑ improving the quality of matches; and,
- ❑ refining the feature extraction process so it can proceed automatically.

It is only by further developing the feature extraction and patch matching algorithms that a robust system can be developed. If the system developed is not robust then it ceases to be a useful piece of applied research. The next section describes some new improvements to the proposed method of feature extraction and registration which lead to improved results.

6.4 New techniques for improved patch matching

6.4.1 Introduction

So far this chapter has presented a reliable method of registering a pair of images based on feature extraction and matching. This section describes a number of new improvements that can be made to the extraction and matching processes which enhance the accuracy and reliability of the results, as well as overcome some of the problems already encountered. The first aspect which is considered is the lack of automation in the feature extraction process. Different methods are presented below for automatically selecting the parameters used in these processes and therefore solving the problem of

manually choosing the most appropriate parameters. The following section describes how results from all of the different feature extraction algorithms can be combined together to improve the number of matched patches found in a pair of images, and ensure that only the best possible matches are used to register the images. § 6.4.4 explains how the feature extraction and feature matching procedures can be integrated to produce an increased number of high quality matched patches. The final section illustrates how these techniques improve the image registration results.

6.4.2 Automatic parameter selection

The automatic selection of feature extraction algorithm parameters is essential if the proposed registration method is to operate without user intervention. Of the work described so far, it is only the feature extraction stage which is not automatic. Therefore this section provides some ideas on how the algorithms can be automated and still produce good results. There are two problems to consider here: what determines the best combination of parameters, and how can these parameters be selected automatically? To answer the first question, it is suggested that the best combination of parameters should give the greatest number of correct matches. Therefore the quantity of correct matches can be used as a measure to select the best combination of parameters. Two solutions are presented which address the second question: a brute force method, and a database/look-up-table method.

The brute force method of parameter selection is very simple but only applicable to some of the feature extraction algorithms. It works by testing a range of parameters for each of the feature extraction algorithms, and seeing which combination of parameters gives the largest number of correct matches. The problem with this method is that it can be time consuming since the whole matching procedure has to be repeated for each different parameter. This problem is not severe if each feature extraction algorithm is dependent on only one parameter, and if the algorithm can be executed quickly, but if this is not the case, then this method of parameter selection could be far too inefficient and it may be quicker to select the parameters manually.

The feature extraction algorithms which only need one parameter to be set are homogeneous patch extraction, MUM and REGSEG. For each algorithm a range of parameters, which were evenly distributed around the values chosen to produce the results shown in the early part of this chapter, were tested. The number of correct

matches found for each combination are shown in the tables 6.7 to 6.10. Matches for Camargue are shown in black, and matches for Istres are shown in red.

		SPOT homogeneous patch extraction parameter					
		Image	3	4	5	6	7
SAR homogeneous patch extraction parameter	11	Camargue	0	0	0	0	0
		Istres	2	3	3	0	1
	12	Camargue	0	0	0	0	0
		Istres	2	3	3	0	3
	13	Camargue	1	0	0	0	0
		Istres	3	3	3	0	3
	14	Camargue	1	0	0	0	0
		Istres	0	0	0	0	0
	15	Camargue	2	0	0	0	0
		Istres	0	0	0	0	0

Table 6.7 Number of correct matches for SPOT homogeneous patch extraction and SAR homogeneous patch extraction

		SPOT REGSEG parameter					
		Image	1	2	3	4	5
SAR homogeneous patch extraction parameter	11	Camargue	0	1	1	1	1
		Istres	2	3	2	2	0
	12	Camargue	0	1	2	1	1
		Istres	2	3	2	2	2
	13	Camargue	0	2	2	2	1
		Istres	2	3	2	2	3
	14	Camargue	1	2	2	1	1
		Istres	0	0	0	0	0
	15	Camargue	1	1	3	2	1
		Istres	0	0	0	0	0

Table 6.8 Number of correct matches for SPOT REGSEG and SAR homogeneous patch extraction

		SPOT homogeneous patch extraction parameter					
		Image	3	4	5	6	7
SAR MUM parameter	4	Camargue	0	0	0	0	0
		Istres	2	2	2	0	4
	5	Camargue	0	0	0	0	0
		Istres	3	3	3	0	3
	6	Camargue	1	0	1	0	0
		Istres	3	2	4	0	4
	7	Camargue	0	1	0	1	0
		Istres	3	4	3	0	3
	8	Camargue	0	1	2	2	0
		Istres	2	3	2	0	2

Table 6.9 Number of correct matches for SPOT homogeneous patch extraction and SAR MUM

		SPOT REGSEG parameter					
		Image	1	2	3	4	5
SAR MUM parameter	4	Camargue	1	2	3	2	1
		Istres	2	2	3	3	3
	5	Camargue	0	3	3	2	1
		Istres	3	4	4	4	4
	6	Camargue	0	3	4	2	2
		Istres	2	4	4	3	3
	7	Camargue	0	1	2	2	1
		Istres	3	3	3	3	3
	8	Camargue	0	1	3	1	1
		Istres	3	2	2	2	2

Table 6.10 Number of correct matches for SPOT REGSEG and SAR MUM

A number of conclusions can be drawn from the results in the above tables. Firstly, and most obviously, the quantity of correct matches found is dependent upon which images are used. Secondly, and more importantly, there seems to be no regular pattern to the results. It is very difficult to draw any correlation between the value of the parameters, and the number of matches found. This is not necessarily a surprising result, but it does mean that a brute force method such as this for selecting the values of the parameters is the only viable automatic method. It is also clear that a number of different combinations of parameter values give the same number of matches, showing that no one particular combination is better than another. Lastly, it appears that the combination of the MUM and REGSEG algorithms give consistently more patches than any of the other three combinations. It is unclear why this should be so, but it is possibly due to the

fact that both algorithm extract features using a similar method, and therefore similar features are extracted from each image.

The results clearly illustrate that this brute force method of parameter selection has worked: combinations of parameters can be selected automatically which will produce enough matched patches to perform the approximate registration. Therefore, the whole registration system so far can be considered automatic. The next section goes on to describe how the large number of extracted patches can be used to improve the accuracy of the registration, but before this, another method of parameter selection is introduced.

The second method of parameter selection described here has not been tested, but would make an interesting topic for future research. The method relies on the construction of a database containing information about different landcover types and the optimal parameters to use to extract features from images of those landcover types. The database would also have to contain information about which sensor acquired the image and how the conditions at the time of image acquisition affect the visibility of particular features. The method would operate by the user giving information about the landcover in the images being registered, and the database providing the optimal parameters for the feature extraction algorithms to give the best results for that particular landcover. Whether or not this method would work in practice is a question that is beyond the scope of this thesis.

6.4.3 Combined use of multiple feature extraction results

One of the strong points of the proposed method of registration, and the theme that runs throughout this thesis, is redundancy. The use of multiple feature extraction algorithms means that it is possible to increase the number of matches found between two images compared to if only one combination of feature extraction algorithms were used. Table 6.5 above compared the number of good matches found between pairs of images, but it does not give any information on whether there is any overlap between matches. For example, if a combination of MUM and REGSEG algorithms gives four matches, and a combination of RWSEG and OPTISEG algorithms gives five matches, then there is a minimum of five matches found between the images, and a possible maximum of nine matches. Thus, by combining the results of these combinations of matching algorithms it may be possible to increase the number of matches, compared to if only one combination were used. Furthermore, where there is overlap of matched features (i.e. where the same match has been found by different combinations of algorithms) it is

possible to select the highest quality match and reject the lowest quality match, thereby improving the accuracy of the matching results.

The results generated from the Istres and Camargue data sets in § 6.3.3 were both used to test the first step of the refinement process: increasing the number of matches by combining matches from different combinations of feature extraction algorithms. This was done by combining all the patches matched from the SAR image into one group, and combining all the patches matched from the SPOT image into another group. By comparing the centroids of the matched patches within each group it is possible to determine where the overlap in the results lies. For the Camargue data set there were a total of 18 matched patches, and for the Istres data set there were 38 matched patches in total. The Camargue patches have been listed in table 6.11.

Reference	SAR grey level	SPOT grey level	Value of cost function	Percentage overlap	Centroid (row, column)	
1	2	4	0.559	72.3	36.665	78.878
2	382	9383	0.907	75.5	177.920	269.584
3	2	1	0.636	80.0	36.665	78.878
4	88	1361	0.308	71.1	386.693	351.594
5	2	4	0.356	73.1	36.821	78.863
6	21	3810	0.186	73.6	398.245	125.198
7	53	9383	0.343	86.0	176.569	267.232
8	88	15579	0.283	78.3	386.693	351.594
9	2	1	0.363	80.2	36.821	78.863
10	64	206	0.641	75.7	177.680	268.424
11	2	4	0.468	71.0	39.268	81.427
12	64	9383	0.193	86.1	177.680	268.424
13	2	1	0.486	74.6	39.268	81.427
14	1072	206	0.315	75.0	177.429	269.559
15	1	4	0.585	73.4	37.105	79.673
16	1072	9383	0.336	86.3	177.429	269.559
17	1	1	0.606	79.5	37.105	79.673
18	136	106	0.362	71.2	47.867	395.934

Table 6.11 Matched patches from Camargue images

Those SAR patches which have centroids within a radius of 10 pixels of each other can be considered to represent the same feature in the SAR image. Thus, line numbers 1, 3, 5, 9, 11, 13, 15, 17 in the above table represent one patch, 2, 7, 10, 12, 14, 16 another patch, 4, 8 another patch, 6 another patch, and 18 another patch. Therefore, there are 5 distinct matched patches. This is greater than the maximum number of matches found before, which was 4.

These matches can now be refined so only the highest quality matches are used in the registration procedure. To determine the best match both the value of the cost function and the percentage overlap must be taken into account. A low cost function and a high overlap is required, so by maximizing the ratio of overlap to cost function, the best results will be highlighted. Table 6.12 shows the details of similar matches grouped together with the ratio of overlap to cost function in the last column. For each group the best match, determined from this ratio, is highlighted. These best matches are displayed graphically in figure 6.35.

Ref.	SAR grey level	SPOT grey level	Value of cost function	%age overlap	Centroid (row, column)		Overlap /cost function
1	2	4	0.559	72.3	36.665	78.878	1.29
3	2	1	0.636	80.0	36.665	78.878	1.25
5	2	4	0.356	73.1	36.821	78.863	2.05
9	2	1	0.363	80.2	36.821	78.863	2.20
15	1	4	0.585	73.4	37.105	79.673	1.25
17	1	1	0.606	79.5	37.105	79.673	1.31
11	2	4	0.468	71.0	39.268	81.427	1.51
13	2	1	0.486	74.6	39.268	81.427	1.53
18	136	106	0.362	71.2	47.867	395.934	1.96
7	53	9383	0.343	86.0	176.569	267.232	2.50
14	1072	206	0.315	75.0	177.429	269.559	2.38
16	1072	9383	0.336	86.3	177.429	269.559	2.56
10	64	206	0.641	75.7	177.680	268.424	1.18
12	64	9383	0.193	86.1	177.680	268.424	4.46
2	382	9383	0.907	75.5	177.920	269.584	0.83
4	88	1361	0.308	71.1	386.693	351.594	2.30
8	88	15579	0.283	78.3	386.693	351.594	2.76
6	21	3810	0.186	73.6	398.245	125.198	3.95

Table 6.12 Best matched patches from Camargue images

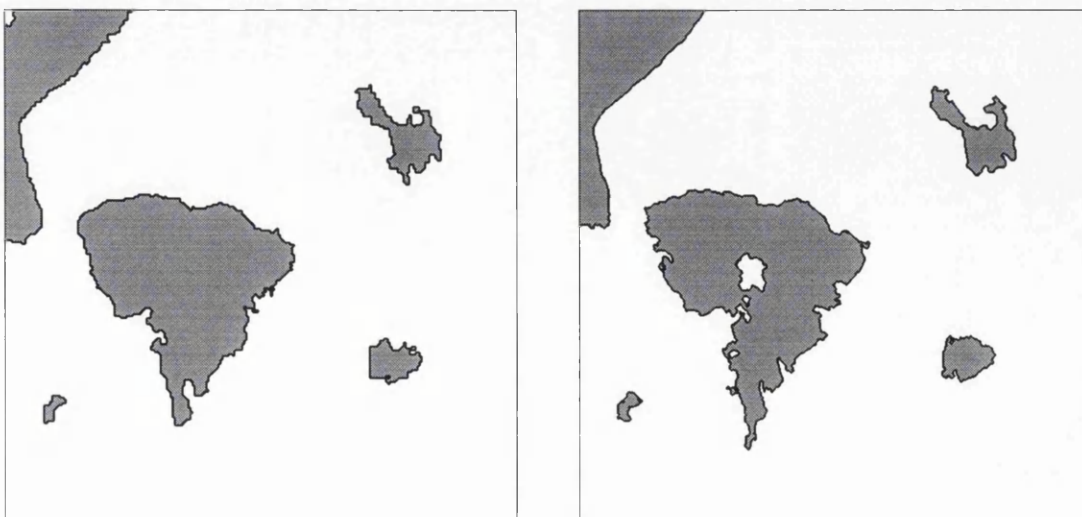


Figure 6.35 Best matched patches from SAR and SPOT images of Camargue (left and right respectively)

The refinement of the matching procedure was repeated for the Istres images. The results are shown graphically in figure 6.36.

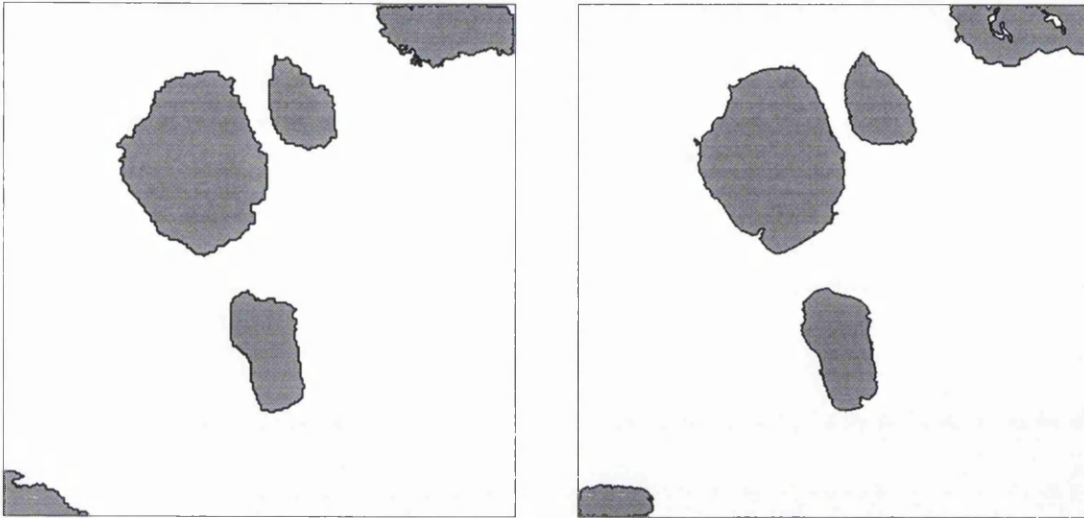


Figure 6.36 Best matched patches from SAR and SPOT images of Istres (left and right respectively)

The value of combining results from different feature extraction algorithms is immediately obvious from the above results. For the Camargue images there are now five matched patches rather than four, and for the Istres images there are five as before, but now the accuracy of the matches is much better. This is best illustrated by comparing figure 6.36 with figure 6.34 which shows the five matches found when the combination of feature extraction algorithms is RWSEG and homogeneous patch extraction. The matched patches in figure 6.36 are obviously very similar to each other, but the matched patches in figure 6.34 are much less similar. Since the latter results are more similar, it is likely that the patches represent the features on the ground more accurately, and therefore they will give better image registration results. Hence, the combined use of different feature extraction algorithms in this way can be considered a success.

The final point to consider in this section is the combination of the previous two concepts. § 6.4.2 described how the parameters of the feature extraction algorithms can be automatically selected by a brute force method, and this section described how the matching results can be improved by selecting patches extracted using different algorithms. It is now proposed that the matching results can be improved still further if all the patches extracted using every possible combination of parameters are matched and refined. This was done for both the Istres and Camargue images, and improved results were obtained.

All the SAR patches extracted in § 6.4.2 were matched with all the extracted SPOT patches. The results were refined and false matches were removed, using the methods described above, until only the best correct matches were left. Other matched patches which were not found here, but were found in the matching performed in § 6.3.3, were combined with these matches.⁴ The results are shown in table 6.13 and 6.14, and displayed graphically in figures 6.37 and 6.38.

Ref.	SAR grey value	SPOT index	Value of cost function	%age overlap	Centroid (row, column)		Overlap /cost function
1	2	2	0.330	76.9	37.743	75.909	2.33
2	112	9383	0.324	85.6	176.441	267.346	2.64
3	1308	615	0.306	75.1	311.012	387.414	2.45
4	78	1584	0.130	75.0	386.662	351.204	5.76
5	22	3810	0.155	75.4	398.301	125.094	4.86

Table 6.13 Best matching results for Camargue image

Ref.	SAR grey value	SPOT index	Value of cost function	%age overlap	Centroid (row, column)		Overlap /cost function
1	12	52	0.116	96.0	191.598	159.453	8.27
2	45	333	0.162	93.0	263.892	345.084	5.74
3	15	123	0.088	88.8	295.350	103.798	10.09
4	86	1390	0.237	71.4	320.588	333.588	3.01
5	2	220	0.115	84.4	444.251	26.058	7.33

Table 6.14 Best matching results for Istres image

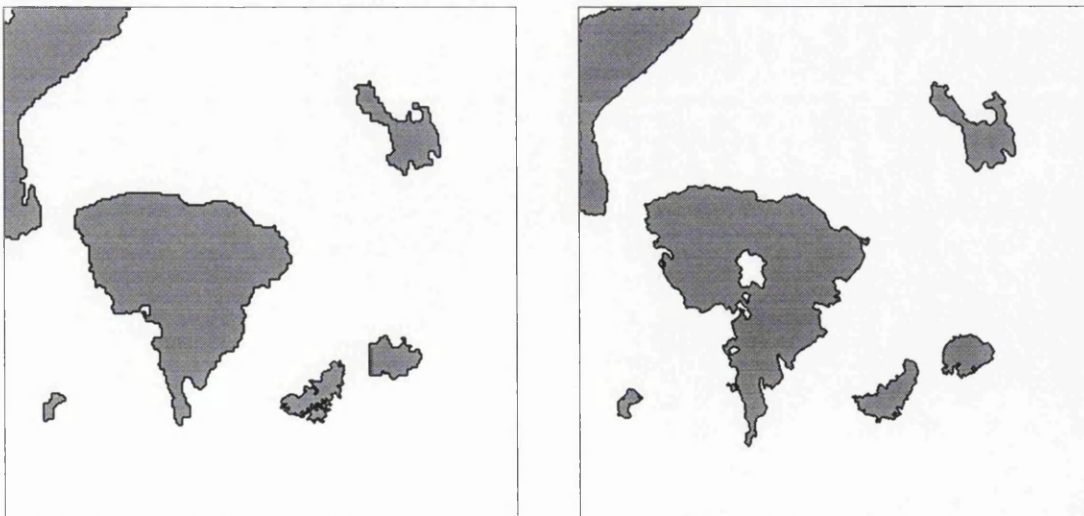


Figure 6.37 Best matched patches from SAR (left) and SPOT (right) images of Camargue

⁴ Remember that § 6.3 utilised all six feature extraction algorithms, whereas this section has only considered four feature extraction algorithms.

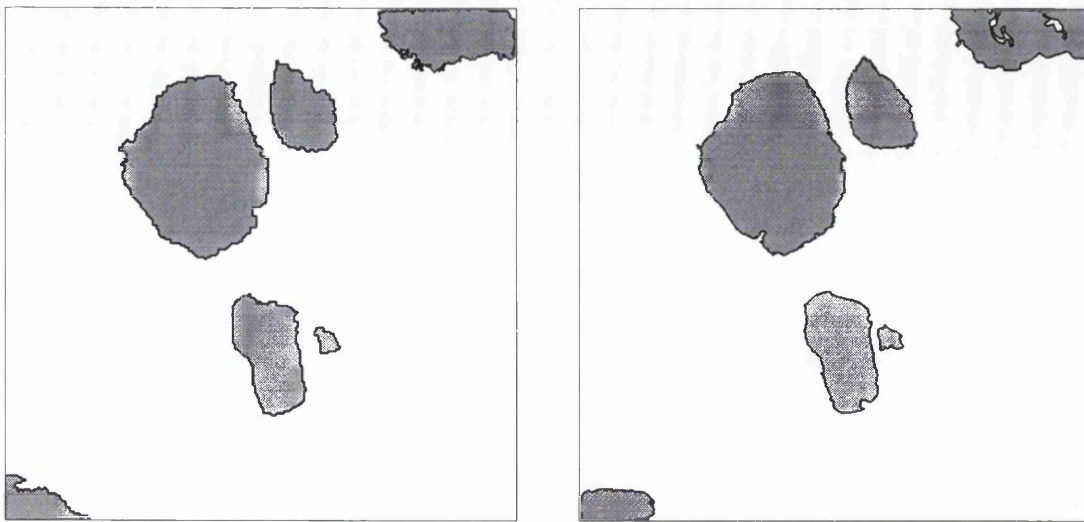


Figure 6.38 Best matched patches from SAR (left) and SPOT (right) images of Istres

This further development to the matching procedure has led to an extra pair of matched patches being found in each of the images, so now there is a total of six matches for each image. The quality of the matches is indicated by the high ratio of percentage overlap to cost function.

This section has illustrated how the matching procedure can be improved in terms of number of correct matches found, and quality of those matches, by combining the results from many different feature extraction algorithms. Results show that the new techniques introduced here are successful for both the test images, and computation times are not prohibitive.

6.4.4 Integrated feature extraction and matching

The idea of combining feature extraction algorithms and feature matching algorithms has already been mentioned in previous chapters, but in this section the concept is developed more fully. The idea is that features are extracted from each image and matched. These matches are then used to approximately register the images as usual, but afterwards the information contained in one image is used to refine the extraction of features from the other image. The matching process is then repeated with the refined extracted features and a better result is obtained. A method of image registration similar to this is described by Li (1995).

Of all the feature extraction algorithms discussed in chapter 5, only the REGSEG algorithm can incorporate information from sources other than the image itself. When used in this way, REGSEG segments optical images by the usual method of region growing, but the growing is limited by edge information provided from another source. In this case that edge information comes from the corresponding SAR image which has

been approximately registered to the optical image. This is summarized in the flow chart in figure 6.39 below. The beauty of this method is that it refines the feature extraction results by making full use of all the information available.

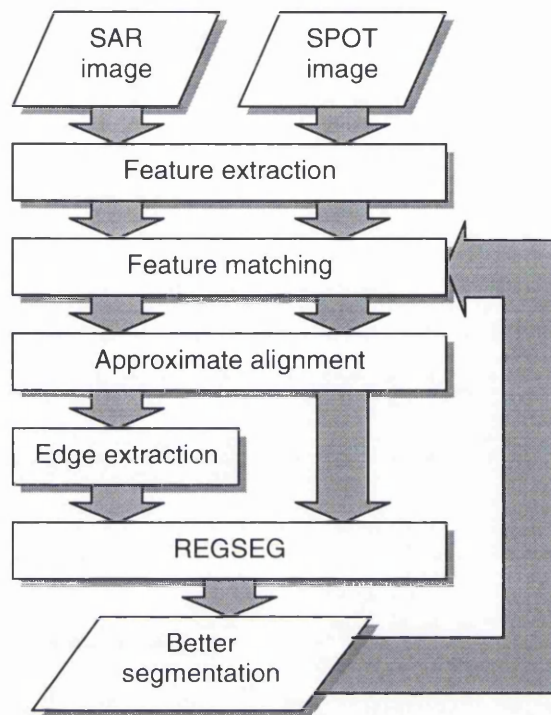


Figure 6.39 Flow chart summarizing integrated feature extraction and matching

This process was tested using the combinations of feature extraction algorithms found in § 6.3.3 which gave the most matches to determine the initial registration. Therefore, for the Camargue test images, the MUM and REGSEG algorithms were used to determine the initial registration, whereas for the Istres test images, RWSEG and homogeneous patch extraction were used. After the images had been approximately aligned, edge images were created from the SAR images by converting the segmented images to edge images. The SPOT images were then re-segmented using this additional information and the matching and registration procedures were followed as before. The results of the refined matching are given in table 6.15 below. The values in the table represent the number of correct matches found when the SAR patches are matched with the refined SPOT patches, created using a range of edge extraction techniques. The numbers in black represent the Camargue data sets, and the numbers in red the Istres data sets.

		Method of SAR patch extraction			
		Auto thresholding	HPE	MUM	RWSEG
Method of edge extraction	None	2 3	2 2	4 4	2 3
	Automatic thresholding	1 3	2 3	2 3	1 3
	HPE	0 3	0 3	0 2	0 2
	MUM	1 3	2 4	1 4	1 4
	RWSEG	1 3	1 3	1 3	1 5

Table 6.15 Results of combined feature extraction and matching

The results in table 6.15 show that in certain circumstances the use of edge information can increase the number of matches found compared to when no edge information is used. However, the success of this method of improving feature matching is dependent upon the combination of feature extraction algorithms used, and the images to which they are applied. The results above show specifically that greater success has been achieved with the Istres image than with the Camargue, but no one method of edge extraction stands out from the others. An important result which is not clear from looking at the figures in table 6.15 is that of the two matches found for the Camargue images using edges derived from automatic thresholding and patches extracted using homogeneous patch extraction, one of them is a new patch. Therefore, when these results are combined with the results from the previous sections, the total number of matches found in the pair of Camargue images has increased to seven.

6.4.5 Results of refined patch matching

The above three sections have illustrated different methods of improving and refining the patch matching results. This section now explains how these improvements are combined with each other, and presents the new registration results which can be obtained. Figure 6.40 shows a flow chart which summarizes the refined matching process.

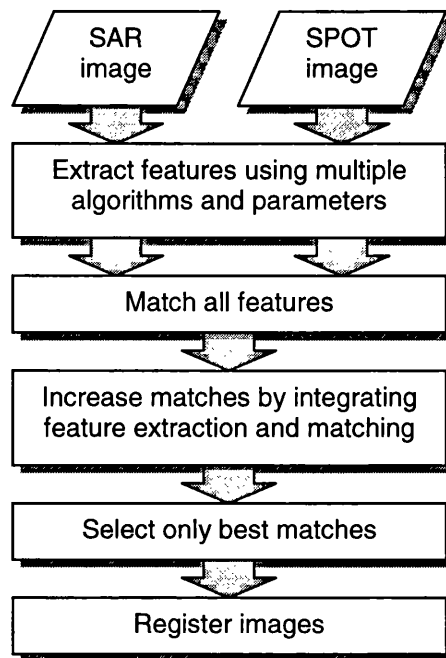


Figure 6.40 Flow chart summarizing image matching with refinements

The first stage of the refined matching process is to extract features from images using all the different feature extraction algorithms available, and as wide a range of parameters as possible. All of the patches extracted are then grouped together into two groups: one for SAR patches and one for SPOT patches. The patches are then matched with each other, and a number of good matches are found. These matches are used to approximately align the images so that the quantity of matches can be increased using integrated feature extraction and feature matching. Any new matches found in this final stage are added to the matches already found. Finally all the matches are used to generate the parameters of an affine transformation to approximately align the images.

The final results of the patch extraction for the SAR and SPOT images of Camargue and Istres are shown in figures 6.41 and 6.42 below. The details of how the patches were extracted are given in tables 6.16 and 6.17.

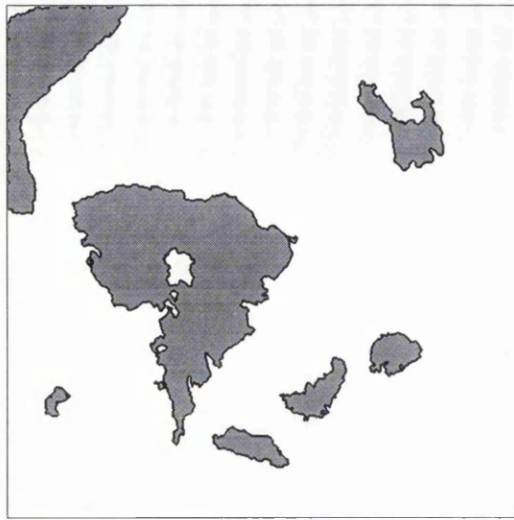
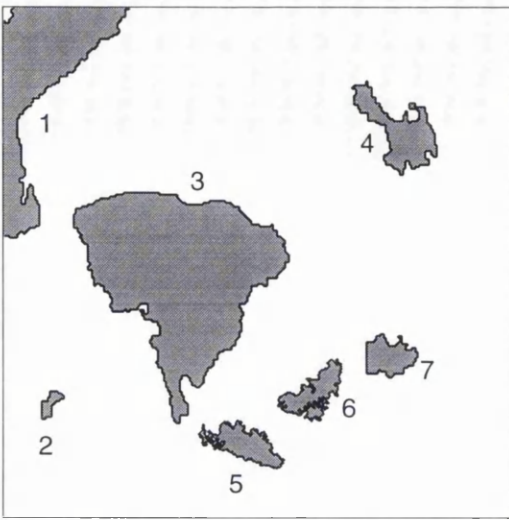


Figure 6.41 Final matched patches from SAR (left) and SPOT (right) images of Camargue

Patch number	SAR		SPOT		Quality of match
	Extraction algorithm	Parameters	Extraction algorithm	Parameters	
1	MUM	$p = 8$	REGSEG	$d = 2$	2.33
2	MUM	$p = 6$	Auto. thresh.	$t = 68$	1.96
3	MUM	$p = 4$	REGSEG	$d = 3$	2.64
4	MUM	$p = 8$	REGSEG	$d = 3$	4.86
5	Hom. patch	$t = 13$	Auto. thresh.	$t = 68$	0.89
6	Hom. patch	$t = 13$	Hom. patch	$t = 3$	2.45
7	MUM	$p = 7$	Hom. patch	$t = 6$	5.76

Table 6.16 Details of matched patches for Camargue images

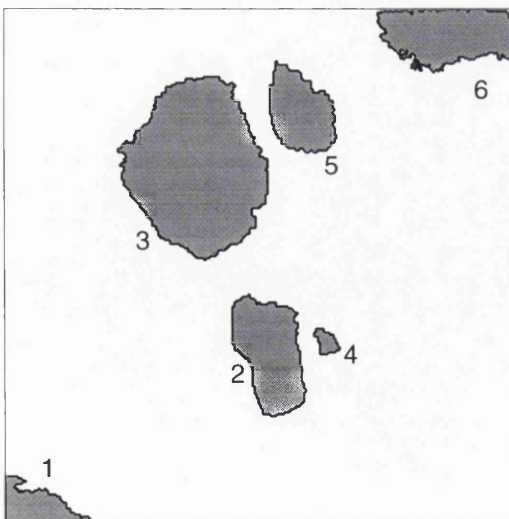


Figure 6.42 Final matched patches from SAR (left) and SPOT (right) images of Istres

Patch number	SAR		SPOT		Quality of match
	Extraction algorithm	Parameters	Extraction algorithm	Parameters	
1	RWSEG	e = 3.29	Hom. patch	t = 3	1.73
2	MUM	p = 6	Hom. patch	t = 3	5.74
3	MUM	p = 6	Hom. patch	t = 3	8.27
4	MUM	p = 4	Hom. patch	t = 7	3.01
5	MUM	p = 5	Hom. patch	t = 5	10.09
6	Hom. patch	t = 13	REGSEG	t = 6	7.33

Table 6.17 Details of matched patches for Istres images

Using the matched patches, the parameters of the affine transformation can be derived for each image. The details of the registration are given below in table 6.18 for Camargue and 6.19 for Istres. As before, the residuals in the tables are the residuals of the tie points and not of independent check points, so therefore they do not give an accurate measure of the quality of the registration. They have been included because they are useful for highlighting the fact that there are no blunders.

	SPOT original		SAR original		SPOT transformed		Residual	
	Column	Row	Column	Row	Column	Row	Column	Row
Camargue	37.83	67.60	37.74	75.91	37.94	68.80	0.12	1.21
	47.87	395.93	46.68	398.84	47.39	396.88	-0.48	0.94
	175.52	266.81	176.44	267.35	176.80	264.70	1.28	-2.11
	399.41	123.88	398.30	125.09	398.21	122.58	-1.20	-1.30
	243.28	441.62	240.91	439.01	241.48	439.74	-1.80	-1.89
	310.11	388.59	311.01	387.41	311.43	388.07	1.32	-0.52
	386.19	348.43	386.66	351.20	386.94	352.10	0.76	3.67

Table 6.18 Co-ordinates of Camargue tie points with residuals

	SPOT original		SAR original		SPOT transformed		Residual	
	Column	Row	Column	Row	Column	Row	Column	Row
Istres	36.30	497.12	29.39	496.42	34.46	494.62	-1.85	-2.49
	263.22	341.24	263.89	345.08	260.36	343.08	-2.86	1.84
	196.22	154.43	191.60	159.45	199.18	156.67	2.96	2.25
	309.51	330.73	320.59	333.59	314.07	331.62	4.56	0.89
	298.92	100.80	295.35	103.80	298.71	100.96	-0.20	0.17
	444.10	25.80	444.25	26.06	441.49	23.16	-2.61	-2.65

Table 6.19 Co-ordinates of Istres tie points with residuals

Patch matching for the Camargue image gave seven tie points, while for the Istres image it gave six. Using these points, the parameters of an affine transformation were generated for each image. The affine transformation function is represented by two equations:

$$X = ax + by + c \quad (6.3)$$

$$Y = dx + ey + f \quad (6.4)$$

The values of the coefficients a to f are listed in table 6.20.

Parameter	Camargue	Istres
a	0.9989	0.9398
b	0.001601	-0.03640
c	9.8839	24.9100
d	0.01059	0.001424
e	1.0156	1.0036
f	-18.6903	-15.629

Table 6.20 Coefficients of affine transformation function for Camargue and Istres

Interesting features to note are that for both images a and e are close to unity, and b and d are close to zero. This signifies that the rotation and scale change between the SAR and SPOT images are small. The coefficients c and f are much larger, signifying that the translation is not small. The affine transformation functions were used to register the SPOT images to the corresponding SAR images. Figure 6.43 shows the outline of the original (untransformed) SPOT features (in cyan), and the outline of the transformed SPOT features (in red) overlaid on top of the original SAR image.

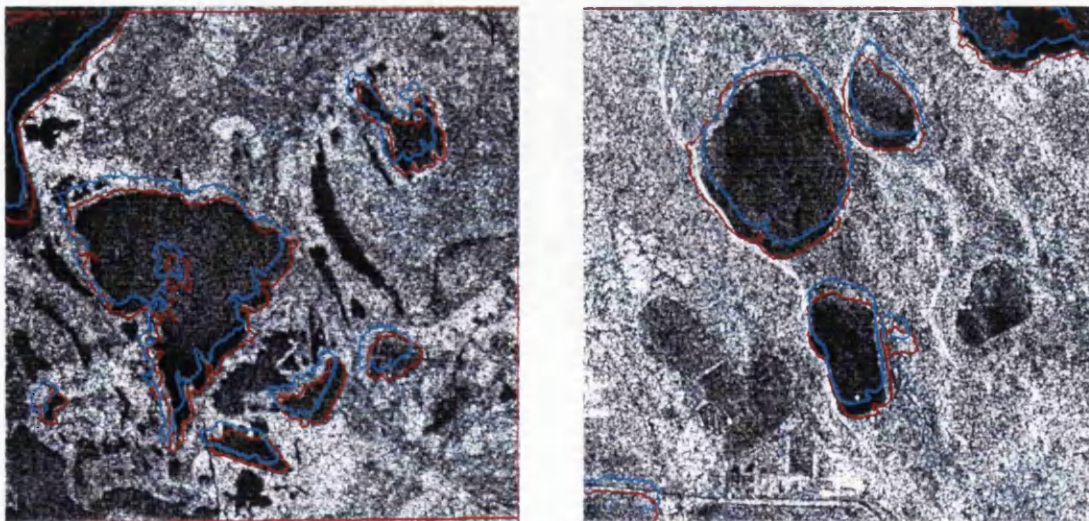


Figure 6.43 Results of patch based registration for Camargue (left) and Istres (right)

It is clear from figure 6.43 that the alignment of the images has been greatly improved. The blue outlines align much more closely with the boundaries of the corresponding

SAR patches than the red outlines. Although the images only provide a qualitative comparison and no quantitative results, they do show that the registration results are sensible and that no blunders have occurred.

A final comment should be made about the distribution of the tie points, since tie point distribution has a marked effect on registration accuracy (Orti, 1981). Due to the automatic nature of the patch extraction and patch matching procedures, it is not possible to control the distribution of the tie points. Therefore, in each pair of images matched using this system, the distribution will be essentially random. A good distribution, where tie points are spread evenly across the image and close to the corners, will not always be obtained, so even if a large number of tie points have been located, the registration could still be quite poor. The only way to increase the chances of a good distribution is to maximise the number of matched patches.

The distribution of tie points located in the pair of Camargue images is slightly more evenly spread across the image than the tie points located in the pair of Istres images. From figure 6.41 it can be seen that the Istres tie points are clustered around a line which runs from top right to bottom left of the image. To ensure a more accurate registration result, a further two tie points would be required in the top left and the bottom right of the image. The Camargue tie points have a more even distribution with points located closer to the corners than in the Istres images (see figure 6.41).

6.4.6 Conclusions

The purpose of this section was to propose refinements that can be made to the standard method of image registration based on patch matching. The reason for doing this was primarily to increase the number of matched patches found between a pair of images, but also to increase the accuracy of those matches. Furthermore, since the results of the matching can be used as a guide for selecting parameters of the feature extraction algorithms, this section also proposed solving the problem of automatic selection of these parameters.

The proposed method of refinement was based mainly on a brute force method: all possible combinations of feature extraction algorithms were used with numerous combinations of parameters. The large redundancy of patches extracted meant that a large number of correct matches were found, and only the most accurate matches were used in the registration. In addition, the brute force method solved the problems of automatic parameter selection. Lastly, a method of integrated feature extraction and

matching was investigated. It was found that the number of correctly matched patches could be increased, and therefore this makes a useful addition to the matching procedure.

The results obtained are clearly better than those generated without the proposed refinements. Without the refined procedure, four correct matches were found for the Camargue images, and five for the Istres images, compared to seven and six matches found when the refined procedure was used, and therefore leading to a better spatial distribution of matches. Also, with the refined procedure the matched patches extracted from each image are more similar to each other, meaning there is a greater likelihood that the centroids actually represent the same point on the ground.

There are many more refinements that can be made to the matching process which have not been discussed here. One important concept is a patch filtering procedure (Dare et al., 1997; Ruskoné and Dowman, 1997). The idea is to filter the extracted patches based on their attributes, so only those patches likely to produce a correct match are actually used in the matching procedure. This technique should increase the efficiency of the whole matching process greatly since the number of patches being matched is reduced. In addition it should also reduce the occurrence of blunders, and therefore increase the reliability of the matching results. It was not tested here since the proposed algorithms have produced enough correct results to prove the concept of alignment of images using extracted patches, but it would make an interesting project for future research.

6.5 Discussion

Although the title of this chapter is simply 'Patch matching' it does in fact describe a large proportion of the image registration procedure. The first section presented the results of applying the feature extraction algorithms described in chapter 5 to the four test images used in this chapter. The parameters of those algorithms were selected manually, and chosen so that the results would provide the best chance of detecting matches between corresponding images.

The following section described the patch matching procedure in detail, from the coding of the patches (i.e. the measurement of the patch attributes), to the matching of the patches, and the calculation of the parameters of the transformation function. The technique of matching patches based on their attributes is not a new one, but few previous studies have applied it specifically to radar and optical satellite data. Due to the very different way in which features on the ground are imaged by these two sensors, this

combination of image types causes problems for the matching process. Although the human observer can clearly distinguish particular features in a pair of images, it is very difficult to extract exactly the same features from each image, and therefore there are few common features to be matched. The consequence is that the matching process must be very robust in order to make best possible use of those common features which have been extracted, and produce the maximum possible number of good matches. The matching procedure proposed in this chapter is obviously reliable and robust since for both pairs of images some correct matches, but no incorrect matches, were found. Furthermore, the matching algorithm was surprisingly time efficient. Although time efficiency is not an important factor in this study (the goal of which is to verify the concept rather than produce an end product) it would affect further development of the proposed techniques into a commercially viable software package.

In previous studies, work often stops at this point: once a suitable number of matches have been found the authors move onto the next subject. However, for the sake of robustness and reliability, this study investigated possibilities for refining the matching procedure in order to increase the quantity and accuracy of the matched patches. A few different techniques were described which lead to a significant increase in the number of correctly matched patches. The most important idea that was proposed is that it is not necessary to rely on just one method for extracting patches: many different techniques can be used and the results combined in order to increase the likelihood of finding correct matches. The consequence was that the number of matches found increased, and the redundancy of similar matches found using different combinations of feature extraction algorithms meant that only the best results need be retained.

The patch matching results were used to register the two pairs of test images. Although there were not enough tie points to allow a rigorous statistical analysis of the residuals, it was still possible to draw some useful conclusions. It was found that the parameters of the transformation function showed that the rotation and scale change between each pair of images was small, but the translation was much larger. This was confirmed when untransformed and transformed boundaries of features from the SPOT image were overlaid on the original SAR image: the translation was obvious, but the rotation and scale change were not noticeable.

This concludes the discussion of patch matching for image registration. This chapter has presented a very successful method of patch matching which includes a number of new

techniques for ensuring a high chance of matching corresponding patches extracted from SAR and SPOT imagery. The alignment of the two pairs of images which have been registered automatically using patch matching is now refined in the next stage of the image registration procedure using edge matching.

Chapter 7 IMPROVING AUTOMATIC REGISTRATION USING EDGE MATCHING

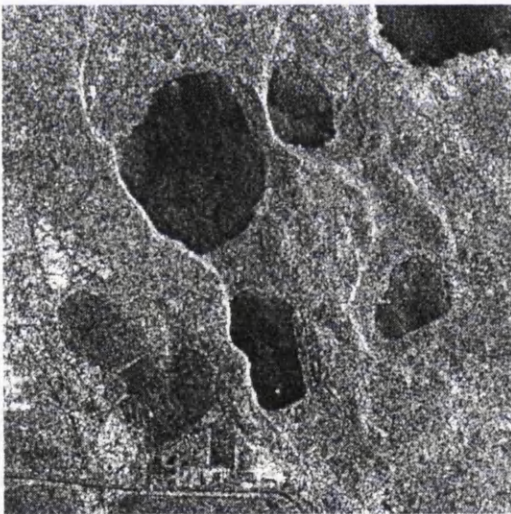
7.1 Introduction

The previous chapter has presented a method of automatically registering images based on patch matching. For some applications which do not require a high level of accuracy of registration, the results produced by this method of automatic registration will be sufficient. However, for most applications, many more tie points will be required to ensure sufficiently accurate registration results. A good spatial distribution of tie points is essential for accurate registration results, as is a large quantity of tie points. In order to improve the quantity and distribution of tie points found in a pair of corresponding images, a refinement stage can be added to the registration process based on edge matching. This chapter presents a description of the application of edge matching to the results generated by patch matching in the previous chapter. It is shown that the number of tie points can be greatly increased without compromising the automatic nature of the proposed registration system. In addition, the increased quantity of tie points enables a fuller statistical analysis of the results, which is included in the latter stages of this chapter.

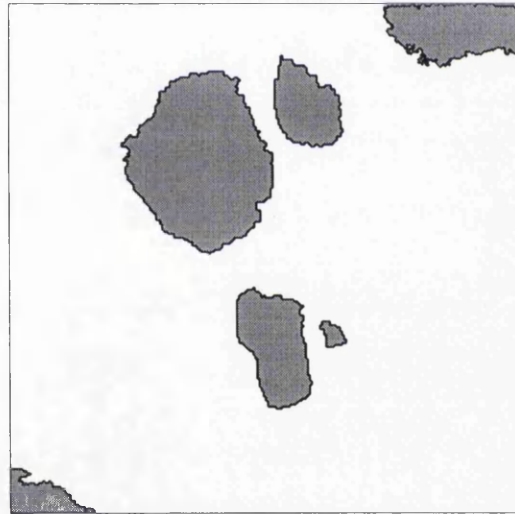
Edge matching can be considered as the final step in the proposed process for automatically detecting tie points between a pair of images. The two reasons cited for using edge matching are (1) to increase the quantity of tie points, and (2) to increase the distribution of the tie points. It is clear this is a necessary step in the processing chain when the results from the previous chapter are considered: patch matching will only produce a few tie points from a pair of corresponding images, even if there are a large number of strong features in those images. However, although edge matching will produce a larger quantity of tie points than patch matching, it cannot be used in isolation: to produce good results, the images have to be approximately aligned in advance, and matched features have to have been located in the pair of images. It is for these reasons that patch matching and edge matching are used together.

Edge matching works by matching pixels which represent the edges of objects on the ground. The edge pixels are extracted using the edge extraction algorithms described in chapter 5 and matched using an algorithm based on dynamic programming.

An important aspect of edge matching, as with patch matching, is that it is essential that only extracted edges which represent the same object on the ground are matched. One way to ensure this would be to compare attributes of the edge being matched (such as length, radius of curvature, location of end points etc.). However, this would involve a lot of processing and would not necessarily produce accurate results. Therefore, the method proposed here is to match only those pixels in the boundaries of patches which are already known to represent the same object on the ground. This is done by using the boundaries of matched patches as masks to eliminate all other unwanted edge information from the images. The mask is created by extracting the boundary pixels of each matched patch as well as all the pixels within a 10 pixel radius surrounding each boundary pixel (see figure 7.1 below).



(a)



(b)

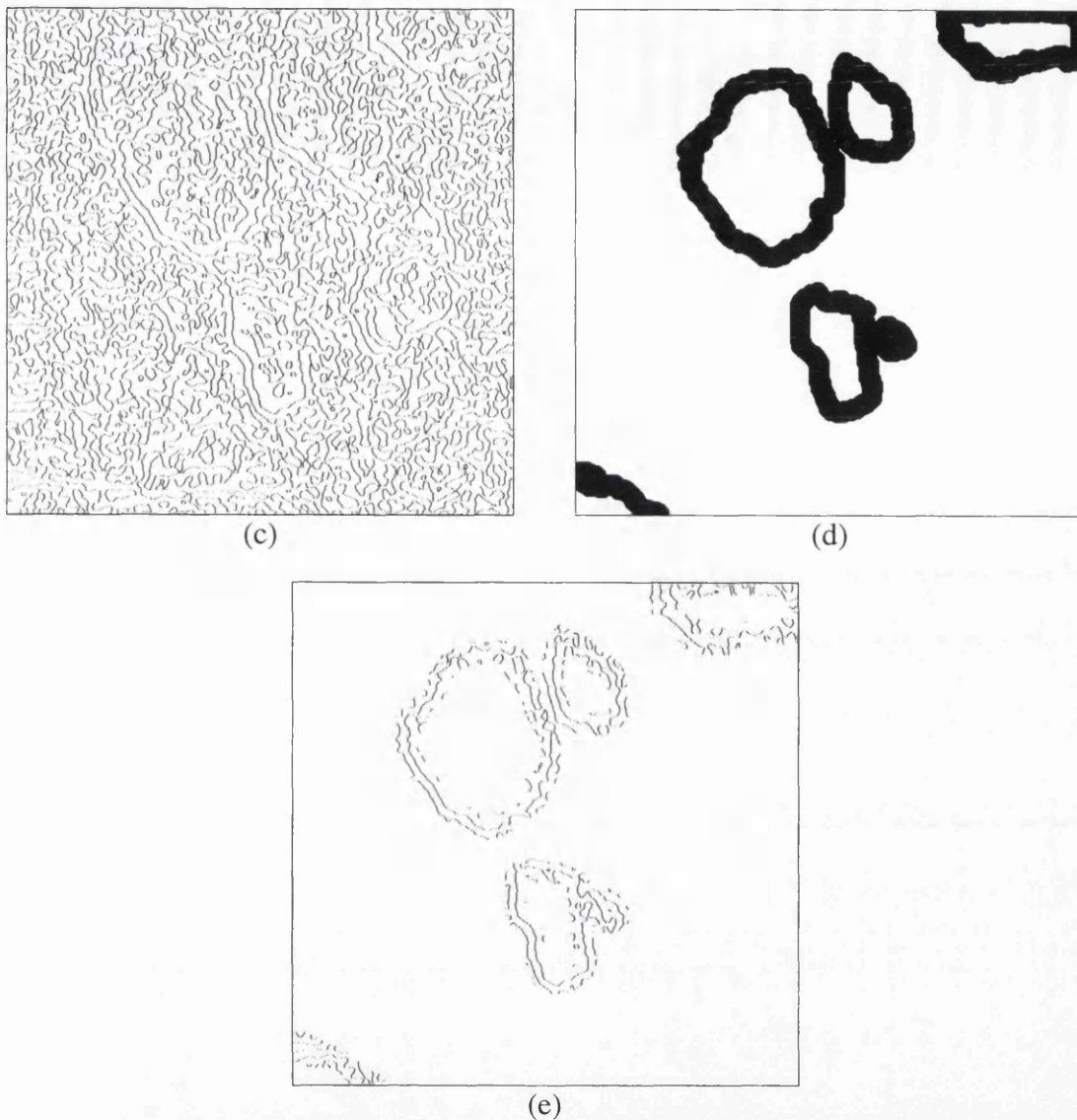


Figure 7.1 Selection of edges for matching using masks

Figure 7.1 (a) shows the unprocessed Istres SAR image, figure 7.1 (b) shows the matched patches and figure 7.1 (c) shows edges extracted from the original image using LoG edge detector. Figure 7.1 (d) shows the mask derived from the matched patches, and figure 7.1 (e) shows the result of applying this mask to the edge image. By comparing figure 7.1 (a) with figure 7.1 (e) it can be seen that a lot of irrelevant edge information has been discarded, and only those edges in the region of the boundaries of the matched features remain. This ensures that only useful edge information will be used in the matching process, leading to a much more efficient algorithm which has a smaller chance of making incorrect matches.

The matching algorithm described here has been used in previous studies (Morgado and Dowman, 1997, Vohra et al. 1996, Newton et al. 1994) and is seen as a reliable method of matching edge features. However, as with the patch matching technique described in the previous chapter, the matching algorithm described in this chapter has been adapted

slightly to increase reliability and accuracy of the results. One significant difference between this study and previous studies is the decision to extract edges by applying edge extraction algorithms to the raw images. Previous studies have in general relied on generating edges by segmenting the image (usually by manual thresholding) and converting the result to an edge image by applying an edge detector to the result. This is an inefficient method of extracting linear features, and is prone to inaccuracies. This study only uses algorithms specifically designed to extract linear features: derivative edge detectors and a SAR edge detection algorithm. It is expected that this will ensure accurate extraction of edges, and therefore reduce the possibility of errors in the matching algorithm.

The first section of this chapter presents the results of applying edge extraction algorithms described in chapter 5 to the test images, whilst the second section describes the matching of the extracted edges and analyses the tie points found.

7.2 Extraction of linear features from SAR and SPOT images

7.2.1 Introduction

This section describes the application of three different linear feature extraction algorithms to the two pairs of test images introduced in the previous chapter: Camargue and Istres. The first and second derivative edge detection algorithms can be applied to both SAR and SPOT data, but the CAESAR edge detection routine is limited to SAR data only. Unfortunately, due to the nature of the edge matching algorithm, only edges extracted using the first derivative edge detector can be used in the matching process. However, results have been generated using the other algorithms so that the first derivative edge detector results can be compared with them.

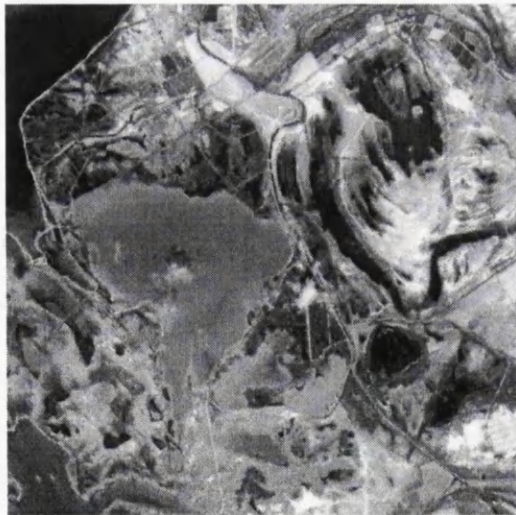
As with the polygon extraction algorithms already described, the results of each edge extraction algorithm depends on which values are used for the different parameters. In order to set the parameters to obtain the best results, it is necessary to define what a good result actually is. The same argument is used here as was used with the polygon extraction algorithms: the best results are those which give the greatest number of good matches. Before the edge features have been matched it is difficult to know what values the parameters should take in order to give the greatest number of matches. Therefore, the parameters used to generate the results in this section were selected by trial and error.

All of the results generated require some degree of post-processing. The principal reason for this is that the edge matching algorithm requires images with edge location and direction information, but no edge strength information. Therefore, the edge strength results generated using the Sobel operator need to be thresholded. However, before thresholding, insignificant edges are removed from the images using masking, and the non-maximal suppression routine described in chapter 5.

7.2.2 First derivative edge detection

The first derivative edge detection algorithm chosen to be used in this thesis is the Sobel edge detector. Although this algorithm can be applied to raw data, preprocessing the data improves the results. Preprocessing of SAR images involves the application of speckle reduction algorithms, whilst preprocessing of optical data is the application of smoothing algorithms.

There are two different smoothing algorithms for SPOT data which have been introduced in this thesis: the median filter and the most homogeneous neighbour filter. The Sobel edge detector was applied to each of the Camargue and Istres images after smoothing with each of these algorithms. Post-processed results are shown in figures 7.2 and 7.3; the threshold level, t , is given in brackets with each image.



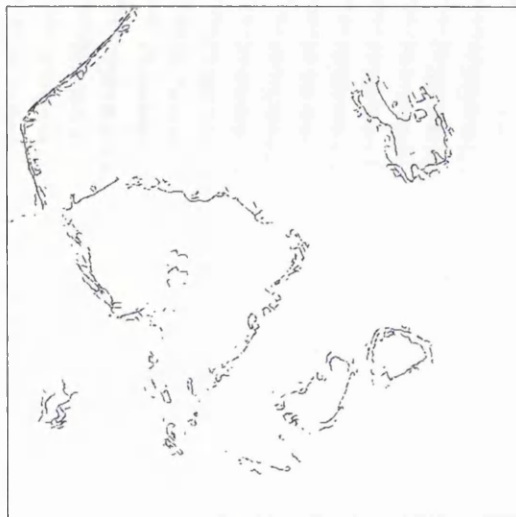
Original image



No smoothing ($t = 40$)

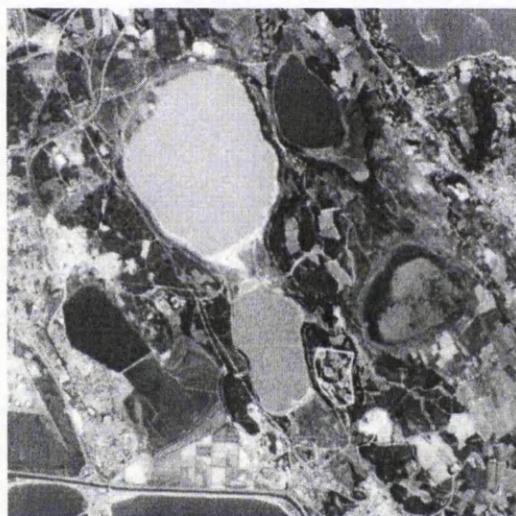


3×3 median filter ($t = 40$)

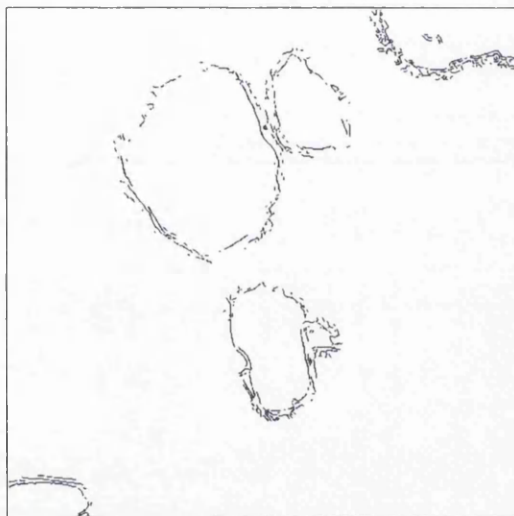


MHN filter ($t = 40$)

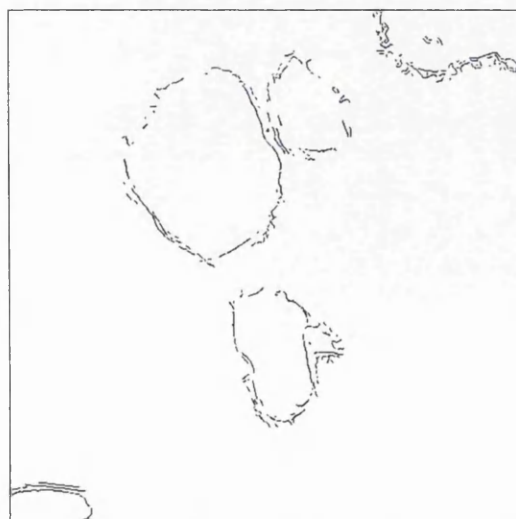
Figure 7.2 Edge images generated from Sobel edge detector applied to smoothed Camargue SPOT images



Original image



No smoothing ($t = 60$)



3×3 median filter ($t = 60$)



MHN filter ($t = 60$)

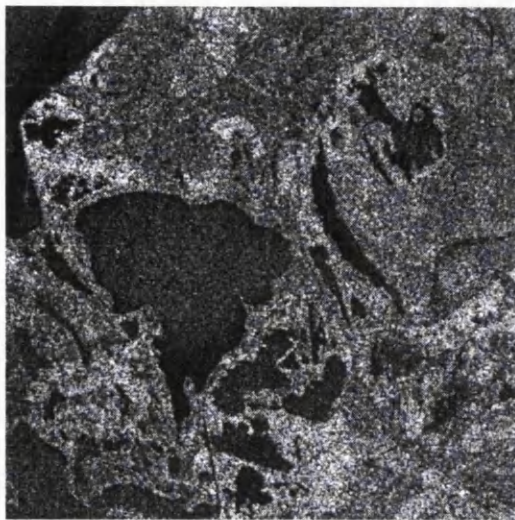
Figure 7.3 Edge images generated from Sobel edge detector applied to smoothed Istres SPOT images

The differences between the three edge images of Camargue and Istres are not great. Although the smoothed images are visually quite different to the unsmoothed images, these differences are not so noticeable when edges are extracted. Even though the results are quite similar, the all three results will be used in the matching algorithm to ascertain if there are any subtle differences making one result better for matching than the others.

Chapter 3 introduced 12 different speckle reduction algorithms for SAR images, all of which gave quite different results. By comparing the results of applying the Sobel edge detector to images processed with each of these algorithms, a decision can be made about which ones are best oriented to aiding the edge extraction procedure. Of the 12 algorithms, the five that were found to produce the most useful results were:

- ❑ median;
- ❑ most homogeneous neighbour (MHN);
- ❑ sigma;
- ❑ sigma median; and
- ❑ MHN median.

The rest of the algorithms were found to cause the Sobel edge detector to extract too much clutter with the edge information and therefore increase the risk of false matches being found during the matching procedure. The fact that the others have been rejected at this stage does not compromise the validity of this study: the five algorithms show that the concept of using multiple feature extraction algorithms is valid, and the results produced are robust. Any future study could incorporate more speckle reduction algorithms to see if the accuracy of the final matching results could be improved. However, in this study, that seems unnecessary. It is interesting to note that two of the five selected filters are new ones which were proposed in chapter 3. This is not necessarily surprising since the goal of the research described in the third chapter was to develop speckle reduction algorithms which could aid the feature extraction process for image registration. Figure 7.4 shows the SAR image of Camargue, and the results of the Sobel filter applied to the five smoothed images. Figure 7.5 shows the same results for the Istres image.



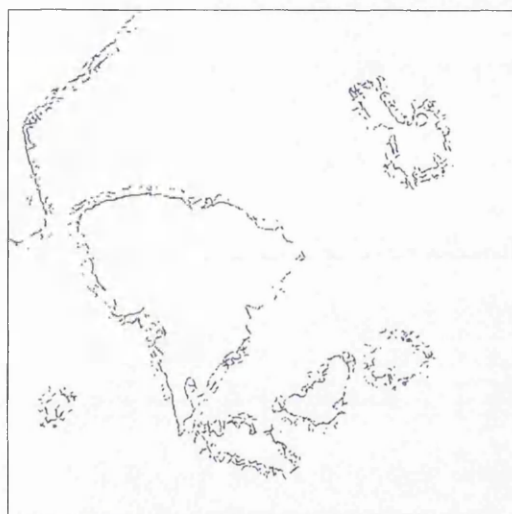
Original image



Median filter ($t = 80$)



MHN filter ($t = 85$)



Sigma filter ($t = 50$)

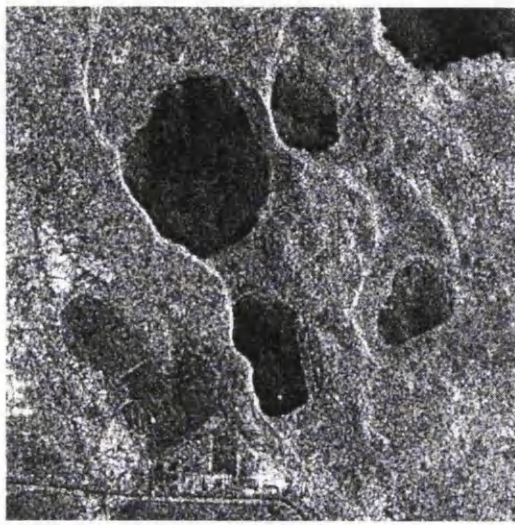


MHN median filter ($t = 70$)

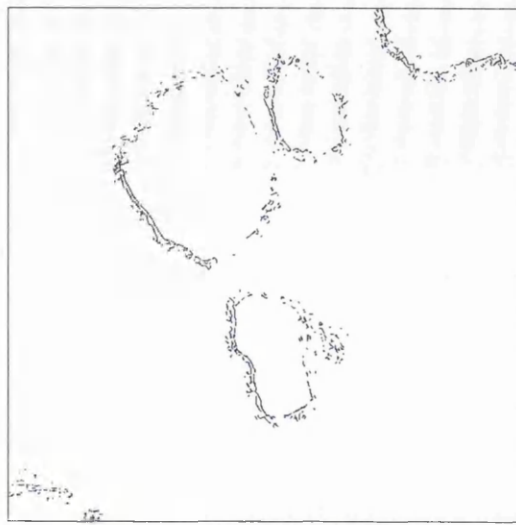


Sigma median filter ($t = 70$)

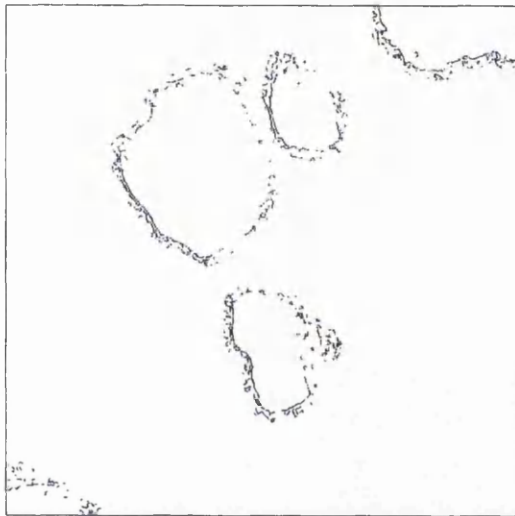
Figure 7.4 Edge images generated from Sobel edge detector applied to Camargue SAR images



Original image



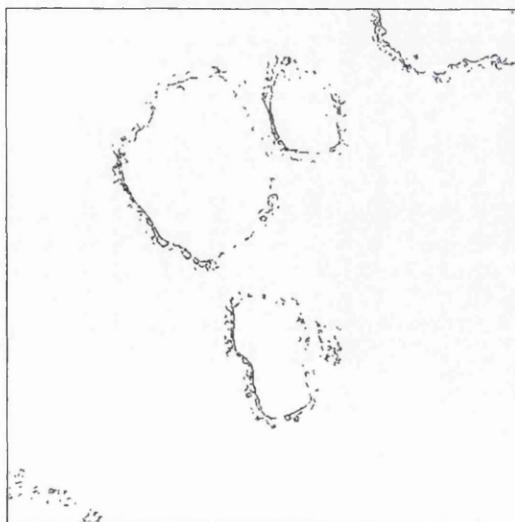
Median filter ($t = 60$)



MHN filter ($t = 65$)



Sigma filter ($t = 35$)



MHN median filter ($t = 60$)



Sigma median filter ($t = 45$)

Figure 7.5 Edge images generated from Sobel edge detector applied to Istres SAR images

The differences between the five results for each of the Camargue and Istres images are slightly more obvious than the differences with the SPOT images, however once again

they are too similar to be able to predict at this stage which one would give the best edge matching results. Therefore, all five results are used in the edge matching procedure.

The full linear feature extraction algorithm using first order edge detection is summarized in figure 7.6 below.

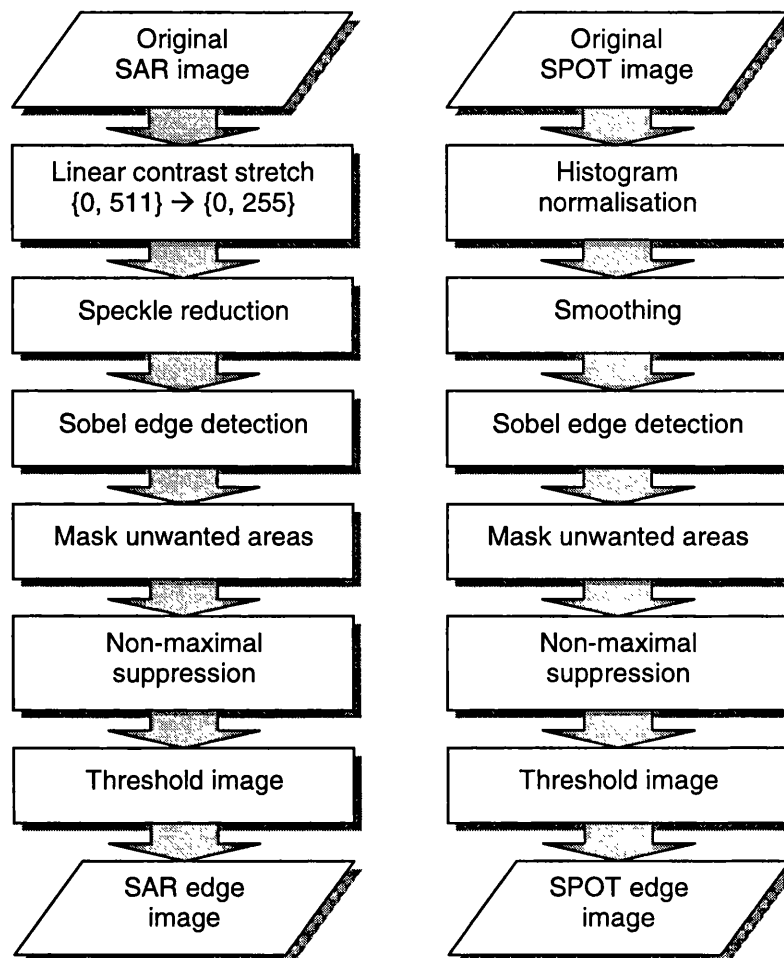


Figure 7.6 Flow chart summarizing linear feature extraction using first derivative edge detection

7.2.3 Second derivative edge detection

The second derivative edge detection algorithm, the Laplacian of Gaussian (LoG), is applied to raw images only. Unlike the Sobel operator, no preprocessing is required since the LoG algorithm incorporates a smoothing component. However, it is necessary to set the value of the parameter w which specifies the width of the window over which the algorithm is applied. A range of values of w were tested for each image, and the results are shown below in figures 7.7 and 7.8. These results have been post-processed by masking out unwanted features.

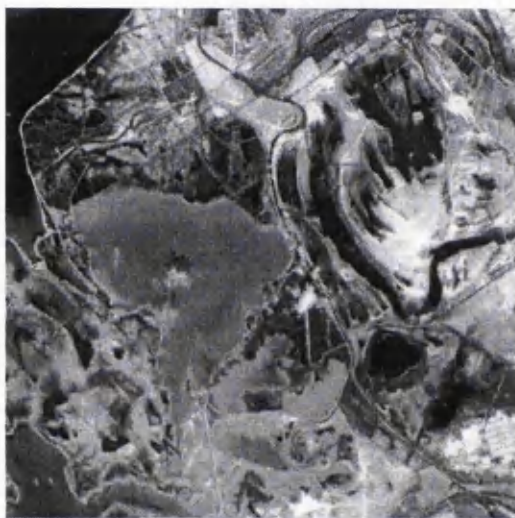
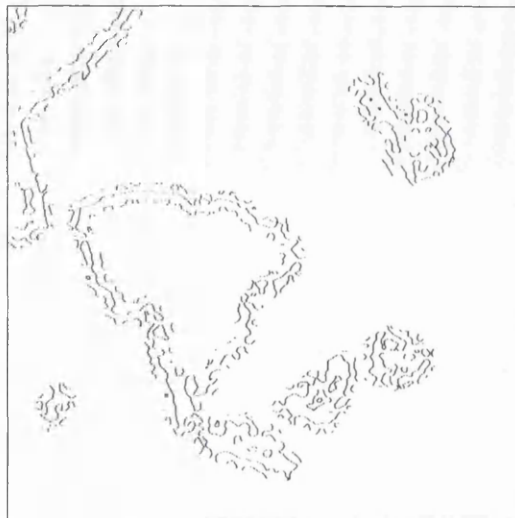
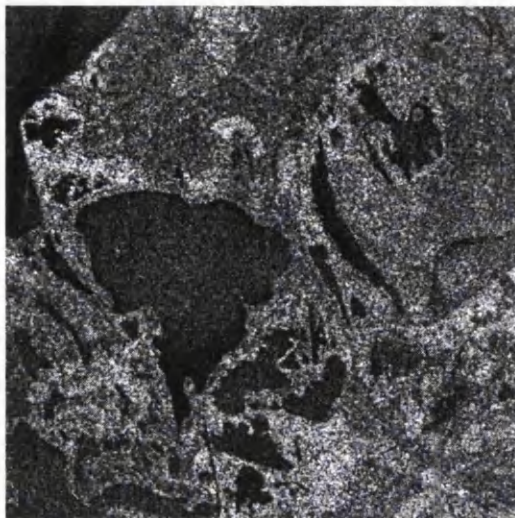


Figure 7.7 LoG algorithm applied to SAR and SPOT images of Camargue

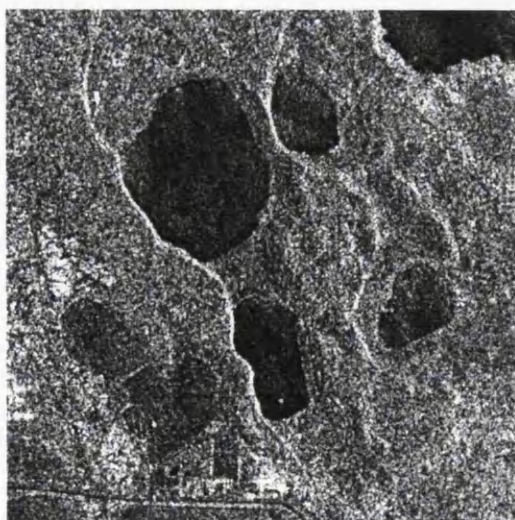




Figure 7.8 LoG algorithm applied to SAR and SPOT images of Istres

The results produced by the Laplacian of Gaussian operator are quite different from those produced using Sobel edge detection. The edges of the features are far more continuous with fewer holes and breaks. However, there is much more ambiguity in the edges: it is more difficult to tell exactly where the edges of the main features are located. The values of w which gave these results were 9 and 10 for the SAR and SPOT images of Camargue, and 9 and 8 for the SAR and SPOT images of Istres. The full algorithm is summarized in figure 7.9.

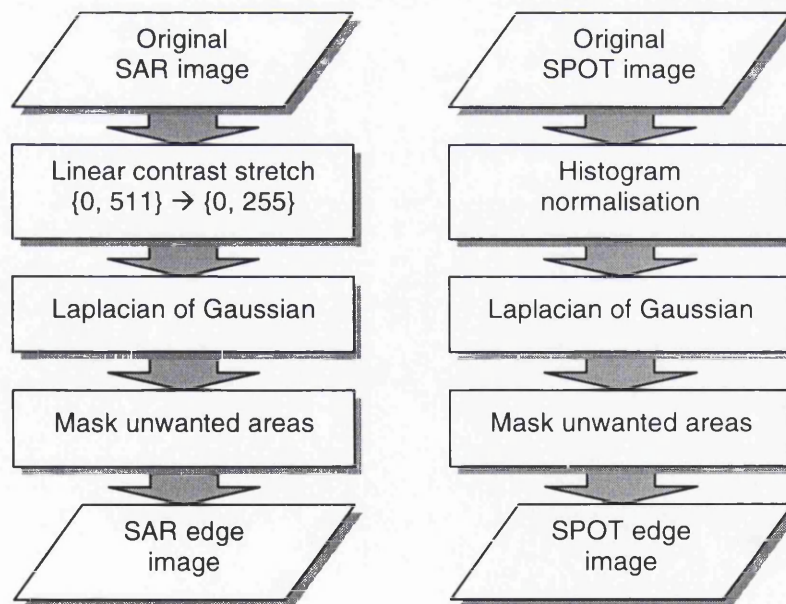


Figure 7.9 Flow chart summarizing the application of the LoG algorithm

7.2.4 CAESAR edge detection

As with the second derivative edge detection algorithm, the CAESAR algorithm is applied to raw image data. A single parameter, p , representing the probability of extracting a false edge has to be selected. For the Camargue image the value of p was

set to 4, and for Istres it was set to 3. The post-processed results are shown in figures 7.10 and 7.11 below.

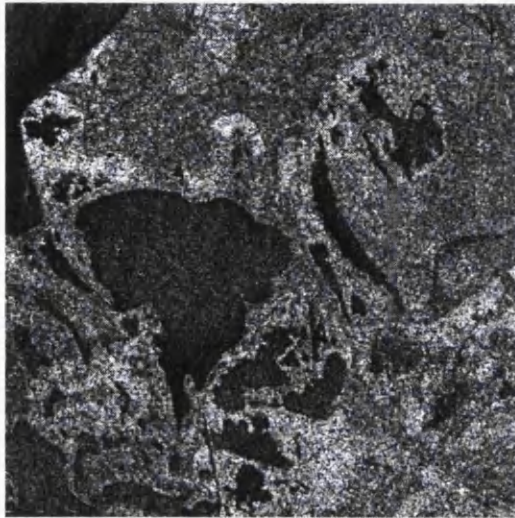


Figure 7.10 Camargue image and result of CAESAR edge detection

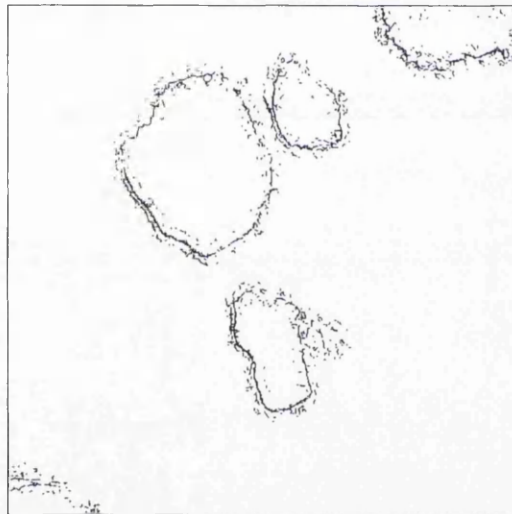
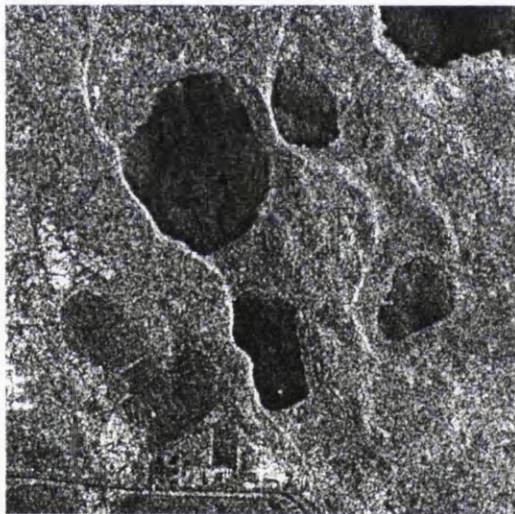


Figure 7.11 Istres image and result of CAESAR edge detection

The CAESAR algorithm appears to produce results which could be very useful for edge matching: the edges are fairly continuous, and unlike the LoG operator there is less ambiguity about the location of the edge. However, some clutter has also been extracted with the edges which may adversely affect the edge matching results. The full algorithm is summarized in figure 7.12.

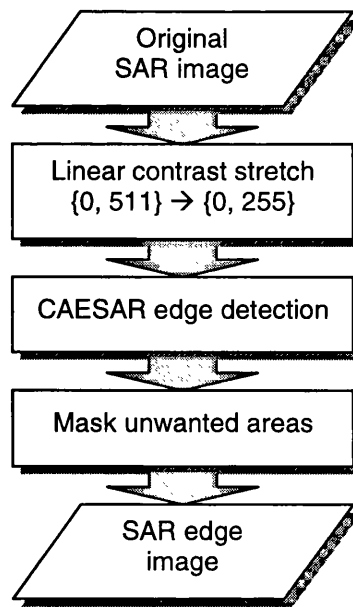


Figure 7.12 Flow chart summarizing the application of the CAESAR algorithm

7.2.5 Conclusions

Each of the edge extraction algorithms has successfully extracted edges from the SAR and SPOT test images, although all of the results are quite different. The CAESAR algorithm appears to be the most successful for the SAR images since it has produced fairly continuous edges which reflect both large and small features in the images. There seems to be few false edges in the resulting images, but a lot of clutter has been extracted with the good edges. If necessary this clutter could easily be removed with post-processing algorithms.

The LoG algorithm produces edge images which are quite different from those produced by the CAESAR or Sobel operators. The edges are smooth and quite continuous with few abrupt changes in direction, but for some features in the image, multiple edges have been extracted, leading to ambiguous results. The results for the SAR and SPOT images are fairly similar.

Each combination of the Sobel operator and smoothing algorithm gives slightly different results, but for the SPOT images these differences are virtually negligible. For the SAR images the differences are still small, but a little more noticeable. The sigma filter apparently produces the most continuous edges, while the results of the other smoothing filters are much the same.

None of the algorithms operate automatically - all of them require some user input. The results generated by the Sobel operator have to be manually thresholded, the LoG operator has a parameter w which has to be set manually, and the CAESAR algorithm

has a parameter p which has to be set manually. Unless a single set of parameters can be found which will produce good results for a range of images, or it is found that the values of these parameters have little effect on the final results, then a method of automatic parameter selection will have to be implemented. This is discussed further in § 7.3.4.

Although the results of the LoG and CAESAR algorithms have been illustrated here, they cannot actually be used in the edge matching process. This is because they produce results which only give information about the location of edge features in the image, and not the direction of those edge features. The matching algorithm relies on both of these pieces of information in order to function properly. The reason for including the results of the LoG and CAESAR algorithms was to enable a comparison of the results of the Sobel operator with results from other sources.

Finally, as a guide for assessing the quality of the edge matching results calculated in the next section, it is useful to know the number of edge pixels extracted using each algorithm. These are detailed in table 7.1 below.

Sensor	Algorithm	Camargue	Istres
SAR	Sobel and median filter	3286	2511
	Sobel and MHN filter	3355	2580
	Sobel and sigma filter	3705	2866
	Sobel and MHN median filter	3583	2267
	Sobel and sigma median filter	3150	2748
SPOT	Sobel and no filter	3990	3065
	Sobel and median filter	3384	2290
	Sobel and MHN filter	3667	2311

Table 7.1 Number of pixels extracted for each edge extraction algorithm

7.3 Edge matching

7.3.1 Introduction

This section describes the edge matching algorithm and presents the results generated from matching the edges extracted in the previous section. The quality of the results are judged according to the number of tie points found, and the accuracy of those tie points. Since it is clear that the edge matching algorithm, in its present form, cannot be

improved or refined, no modifications have been suggested. However, this algorithm is not the final word on edge matching, and suggestions for an altogether more robust algorithm are deferred to the discussion at the end of this chapter.

7.3.2 Edge matching by dynamic programming

The edge matching algorithm (Newton et al., 1994) uses dynamic programming and is loosely based on a method described by Maitre and Wu (1989). For it to produce useful results, an approximate transformation must be defined between the SAR image and the SPOT image. In this case, that transformation is determined from the centroids of the patches which have already been matched. The algorithm operates by moving along each edge in the SAR image pixel by pixel, and by using the transformation function it predicts where each pixel will be located in the SPOT image. For each SAR pixel, a search area is set up around the predicted location in the corresponding SPOT image. A cost function, based on pixel separation and difference in direction, is evaluated for all SPOT pixels within a certain radius of the predicted position of the SAR pixel, and the SPOT pixel with the lowest cost function value is accepted as the match. Only one SPOT pixel can be matched to each SAR pixel, and *vice versa* (see figure 7.13 below).

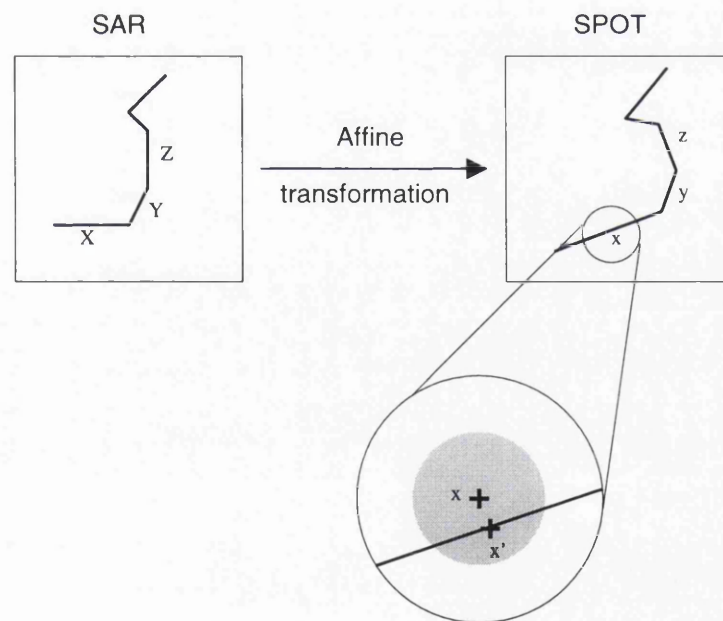


Figure 7.13 Linear feature matching

The predicted location of pixel X in the SAR image is x in the SPOT image. A search is made around point x in the SPOT image (denoted by the grey area) and pixel x' is found to be the lowest cost match. The algorithm then proceeds onto pixels at Y and Z.

The most important aspect of this algorithm is that when all the edges in the SAR image have been considered, the edge with the lowest average cost function is used to update

the initial transformation. The whole matching process is then repeated, but this time it starts from new, and most likely better, starting conditions. However this time, the edge used to update the transformation is excluded from the matching. This process is repeated until all of the edges in the SAR image have been excluded.

7.3.3 Edge matching results

The matching algorithm can only be applied to the edges extracted from corresponding SAR and SPOT images using the first derivative edge detector (Sobel operator). The reason, as described above, is that the algorithm requires both edge location and edge direction information. Unfortunately, of the extraction algorithms detailed here, only the Sobel operator provides this information. Since five speckle reduction filters were tested with the SAR images, and three smoothing options were employed with the SPOT images, for each pair of images there are a possible 15 combinations of edge extraction algorithms.⁵

The matching results (i.e. the number of tie points found for each pair of edge extraction algorithms) are shown in table 7.2 for the Camargue images, and table 7.3 for the Istres images. An obvious trend in table 7.1 above is that more edge pixels have been extracted from the Camargue images than the Istres images. This would imply that more matched edge pixels will be found for the Camargue images than the Istres images. Therefore, in order to be able to compare the results from each image, the ratio of matched pixels to total number of extracted edge pixels for the two images in question has been expressed as a percentage in tables 7.2 and 7.3 below.

⁵ In the following discussion, each combination of Sobel operator and smoothing filter is referred to as an 'edge extraction algorithm.'

		Method of edge extraction from SPOT image		
		Sobel and no smoothing	Sobel and median filter	Sobel and MHN filter
Method of edge extraction from SAR image	Sobel and median filter	1813 (25%)	1623 (24%)	1745 (25%)
	Sobel and MHN filter	1802 (25%)	1593 (24%)	1661 (24%)
	Sobel and Sigma filter	1901 (25%)	1678 (24%)	1938 (26%)
	Sobel and MHN median	1929 (25%)	1692 (24%)	1858 (26%)
	Sobel and sigma median	1740 (24%)	1492 (23%)	1632 (24%)

Table 7.2 Number of tie points found with percentages for Camargue image using different combinations of edge extraction algorithms

		Method of feature extraction from SPOT image		
		Sobel - no smoothing	Sobel - median filter	Sobel - MHN filter
Method of feature extraction from SAR image	Sobel - median filter	1531 (27%)	1120 (23%)	1122 (23%)
	Sobel - MHN filter	1503 (27%)	1116 (23%)	1111 (23%)
	Sobel - sigma filter	1681 (28%)	1248 (24%)	1265 (24%)
	Sobel - MHN median	1400 (26%)	1053 (23%)	1033 (23%)
	Sobel - sigma median	1615 (26%)	1178 (23%)	1150 (23%)

Table 7.3 Number of tie points found with percentages for Istres image using different combinations of edge extraction algorithms

There are no obvious trends in the results shown in tables 7.2 and 7.3, suggesting that the number of tie points found is independent of image and independent of which edge extraction algorithm is used. For all combinations, approximately a quarter of the total

number of extracted pixels in the two images are found to match each other. Although there are slight variations in this figure, those variations are inconsistent, and therefore it is not possible to draw any hard and fast conclusions. In order to further analyse the matching results, the next section considers the accuracy of the matched pixels, rather than the quantity of matched pixels.

7.3.4 Image registration

The results generated in the previous section can now be used to register each pair of images. For all combinations of feature extraction algorithms there are enough tie points to be able to assess the accuracy of the registration. For each case the matched points were randomly split into two equal sized groups: tie points and check points. The tie points were used to generate the parameters of the affine transformation which in turn was used to register the images. Subsequently these parameters were used to transform the check points. Residuals were calculated by comparing the transformed location of the check point with the actual location. Since the check points were not used in the calculation of the transformation parameters, they can be considered independent, and therefore a reliable indicator of the accuracy of the registration. Tables 7.4 and 7.5 show the root mean square (RMS) residual for each combination of edge extraction algorithm for each pair of images.

		Method of feature extraction from SPOT image		
		Sobel - no smoothing	Sobel - median filter	Sobel - MHN filter
Method of feature extraction from SAR image	Sobel - median filter	1.652	1.672	1.772
	Sobel - MHN filter	1.743	1.668	1.737
	Sobel - sigma filter	1.779	1.869	1.858
	Sobel - MHN median	1.635	1.633	1.701
	Sobel - sigma median	1.692	1.670	1.910

Table 7.4 RMS residuals for Camargue image transformation for different combinations of edge extraction algorithms

		Method of feature extraction from SPOT image		
		Sobel - no smoothing	Sobel - median filter	Sobel - MHN filter
Method of feature extraction from SAR image	Sobel - median filter	1.693	1.601	1.650
	Sobel - MHN filter	1.488	1.546	1.662
	Sobel - sigma filter	1.595	1.632	1.624
	Sobel - MHN median	1.597	1.739	1.606
	Sobel - sigma median	1.607	1.534	1.495

Table 7.5 RMS residuals for Istres image transformation edge different combinations of edge extraction algorithms

For all combinations of edge extraction algorithms for both images, the RMS residuals fall in the range of 1.5 to 2.0 pixels. The general trend is that much the same results are obtained for whatever combination of edge extraction algorithms are used. This is not wholly unexpected if the images of the extracted edges (shown in figures 7.2 to 7.5 in § 7.3) are compared. Since the differences between the edge images are small, the RMS residuals generated from those edge images would not be expected to differ greatly.

The two attributes which have been used to assess the quality of the edge matching are the ratio of tie points located to total points extracted, and the RMS residuals of the tie points. In order to gain a more of an insight into the quality of the tie points, two more factors should be considered: the spatial distribution of the tie points and the angular distribution of the residuals. Both of these factors can be displayed simultaneously by plotting the location, magnitude and direction of each residual. Figure 7.14 shows the residuals for two Camargue matching results: (1) SPOT Sobel with median filter and SAR Sobel with MHN median filter, and (2) SPOT Sobel with MHN filter and SAR Sobel with sigma median filter. Similarly figure 7.15 shows the residuals for two Istres matching results: (1) SPOT Sobel with no filter and SAR Sobel with MHN filter, and (2) SPOT Sobel with median filter and SAR Sobel with MHN filter. In each case the two matching results have been chosen since they represent the lowest and highest residuals for each test scene respectively. Note that for each case only 10% of the total number of residuals have been plotted (so that the images are not over-cluttered), and that the magnitudes of the residuals have been increased by a factor of eighty to make

them visible in the images. Although only four results are illustrated here, they are typical examples of all thirty combinations described in tables 7.4 and 7.5 above.

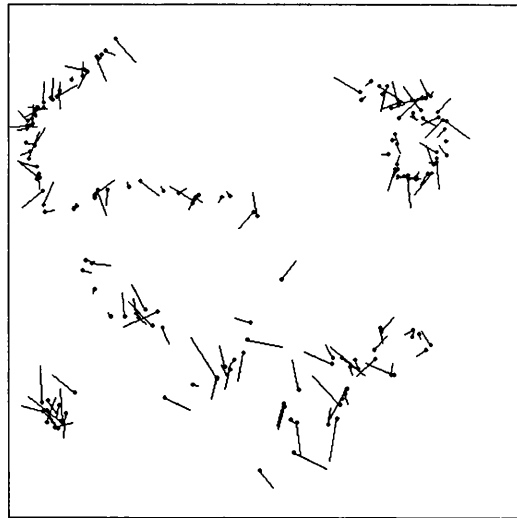
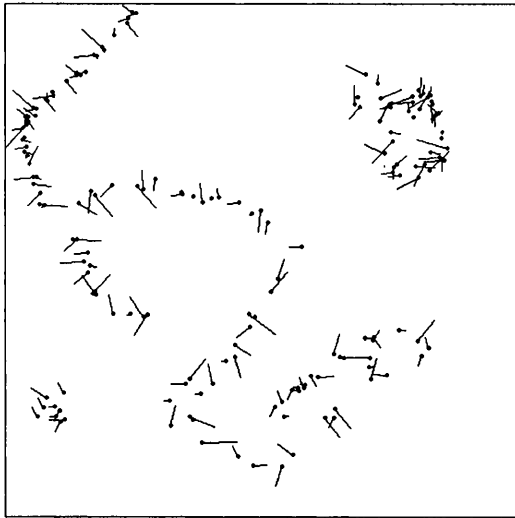


Figure 7.14 Distribution of residuals for two Camargue matching results

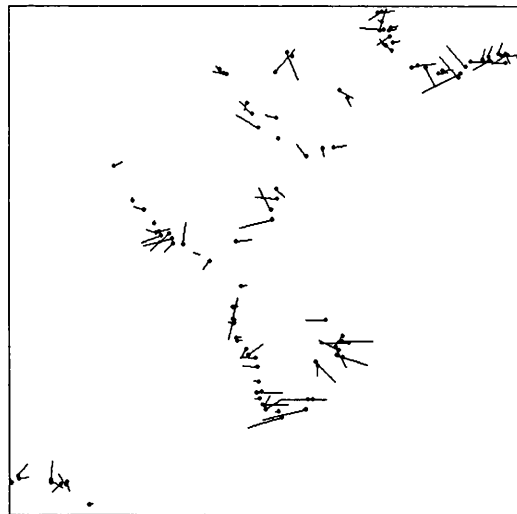
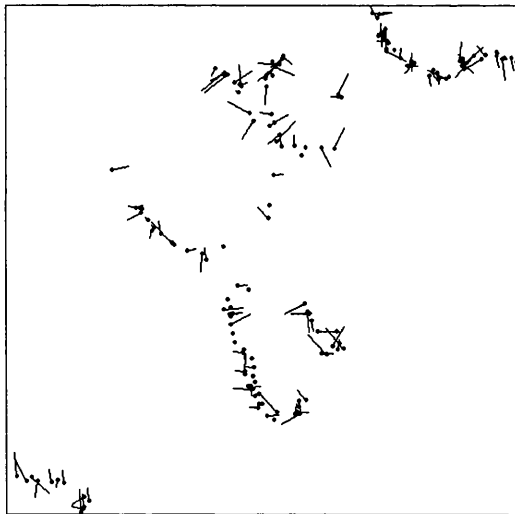


Figure 7.15 Distribution of residuals for two Istres matching results

Since the edge matching procedure matches points in the region of the boundaries of the matched patches, the spatial distribution of the tie points determined from edge matching are going to be similar to the spatial distribution of the tie points determined from patch matching. Even so, the distribution is much improved since the boundaries of the patches lie closer to the edges of the image than the centroids of the patches. It was stated in the previous chapter that the distribution of tie points located in the Camargue images was better than the distribution of tie points located in the Istres images. The same is also true here. Comparing figures 7.14 and 7.15 it is clear that the spatial distribution of the Camargue check points is more evenly spread across the images than the Istres check points. As with the patch matching results, the Istres check

points are again clustered around a diagonal line running from the top right of the image to the bottom left. However, this time that clustering is not so severe.

As well as depicting the spatial distribution of the check points, figure 7.14 and 7.15 also show the angular distribution (directions) of the residuals. A visual examination of the images suggest that the directions are random – there is no reason to suspect systematic errors or blunders are present.

One final method of assessing the quality of the tie points is to visualize the registration results. Since the above matching results have not been able to identify any particular edge extraction combination as being better than any other, just one combination was chosen at random to generate the registration results: Sobel-median with Sobel-median. This combination provided 1623 tie points for the pair of Camargue images and 1120 tie points for the pair of Istres images. These tie points were used to generate the parameters of an affine transformation (see equations 7.1 and 7.2), which are listed in table 7.6 below.

$$X = ax + by + c \quad (7.1)$$

$$Y = dx + ey + f \quad (7.2)$$

Parameter	Camargue	Istres
<i>a</i>	1.0002	0.9998
<i>b</i>	0.0006225	-0.0002092
<i>c</i>	-0.1666	-0.05332
<i>d</i>	-0.0007973	-0.03367
<i>e</i>	1.0002	0.8949
<i>f</i>	-0.1808	0.3010

Table 7.6 Coefficients of affine transformation function for Camargue and Istres

It should be remembered that the edge matching was performed as a refinement technique using the original SAR image and the transformed SPOT image as input images. Therefore, the edge matching affine coefficients are not related to the patch matching coefficients calculated in the previous chapter. However, there are some similarities: once again *a* and *e* are close to unity, and *b* and *d* are close to zero, signifying a small rotation and a scale factor close to one. This time *c* and *f* are small as

well, which is expected since the edge matching algorithm is a refinement to the patch matching results. Edge matching transformation parameters which are large relative to the patch matching transformation parameters would not signify a refinement and therefore imply that something was wrong with one of the matching processes.

The edges of the SPOT features, transformed using edge matching, have been overlaid on top of the SAR image in figure 7.16. The edges of the SPOT features transformed using patch matching have not been displayed in this image since they lie very close to the edges of the SPOT features transformed using edge matching and would confuse the results.

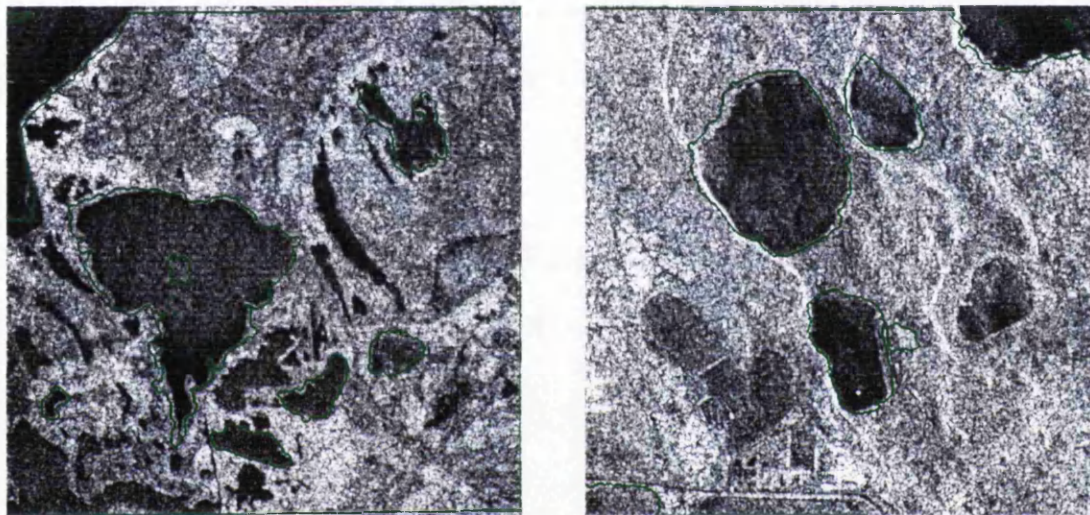


Figure 7.16 Outlines of transformed SPOT features (green) overlaid on SAR image

Figure 7.16 shows that the transformed SPOT edges are aligned well with the edges of the SAR features, confirming that the refinement of the automatic registration procedure has not introduced any new errors or blunders.

In conclusion, it is apparent from the results in tables 7.2 to 7.5 and the subsequent spatial analysis of the residuals that the choice of smoothing algorithms for SAR and SPOT images does not have a great influence on the final edge matching results. It is therefore reasonable to assume that the choice of parameters used in the edge extraction algorithms will also have little effect on the edge matching results. For this reason it is proposed that the parameter values used in these two tests should produce acceptable results when applied to other images. This fact is verified in the next chapter.

7.3.5 Conclusions

This section has introduced a method of finding tie points in a pair of images based on edge matching, and successfully applied it to features extracted from SAR and SPOT

images. The method has been used in previous studies and is accepted as being both reliable and robust. The edge matching algorithm was applied to all the combinations of edge images which were generated using the Sobel operator in section 7.2 and the results were appraised by comparing the number of tie points found, the accuracy of those tie points, the spatial distribution and the angular distribution of the residuals, for each combination.

The accuracy of the tie points was assessed by examining the residuals generated when those tie points were used to calculate the parameters of the transformation function. In order to ensure an unbiased result, one half of the tie points were used in the calculation of the parameters, and the other half were used as check points. The RMS residuals for each combination of edge extraction algorithms were calculated and compared, and it was found that the RMS residuals all fell in the range of 1.5 to 2.0 pixels. Analysis of the distribution of the tie points showed that they were spread fairly evenly across the image, and that the directions of the residuals were apparently random in nature. When the affine transformation was used to register the images, the edges of the principal features in the SPOT data were found to match well with the corresponding features in the SAR data.

The edge matching results show that the choice of smoothing algorithm applied to the image before the Sobel operator is applied does not make any noticeable difference to either the number of tie points found, or the accuracy of those tie points. This fact simplifies the automatic registration procedure since, unlike in the patch matching procedure, results from multiple feature algorithms do not have to be generated.

7.4 Discussion

This chapter has presented a method of refining the registration results generated by the patch matching technique described in the previous chapter. The two reasons for using this refinement technique are (1) to increase the quantity of the tie points, and (2) to increase the distribution of the tie points. To achieve these aims an edge matching technique was employed, and the results generated showed that the technique was successful.

Although edge matching techniques have been used in previous registration studies a few new ideas have been incorporated into this interpretation of the technique. The first new concept is associated with the extraction of the edges themselves. Other studies have used differing methods of extracting edges, but this study proposes a method

which makes better use of the edge extraction algorithms: edges are extracted by applying different operators to raw or preprocessed images, after which all unwanted edge information is discarded using masks generated from the patch matching results of the previous chapter. This method ensures that reliable edges are extracted from the images, and that all unwanted information is removed before any processing takes place.

The next new concept to be proposed is the use of edge images generated by multiple edge extraction algorithms. Unfortunately the rigid nature of the matching algorithm only allows edges generated by first derivative edge detectors to be used, so this valid of this idea could not be tested. However, as a result of this, one final proposal can be made: that the edge algorithm is replaced by a more efficient and reliable one.

Although the algorithm tested in this chapter gave good results, there would be no harm in improving the algorithm to increase its usefulness. The present algorithm matches edge pixels based only on location and edge direction. If pixels in an edge could be matched using more of their attributes, then it seems likely that a more accurate and reliable result would be obtained. Other attributes which could be used are the length of the edge, the radius of curvature of the edge, the location of the pixel within the edge, location of start and end points of the edge, and so on. If this were the case, then the results would not be so dependent on the two attributes which are used in the present algorithm.

This chapter brings to a close the discussion of the development of an automatic registration system. The next chapter aims to apply the techniques developed in this thesis to a pair of full scene SAR and SPOT images in order to accurately register them. A discussion of the techniques developed, and the quality of the results they generate is presented in the final chapter.

Chapter 8 AUTOMATIC REGISTRATION OF FULL SCENE IMAGES

8.1 Introduction

This thesis has proposed a method of automatic image registration based on feature matching. The previous chapters have explained in detail how each of the individual algorithms operate, and results have been demonstrated for a range of test images. This chapter now applies those algorithms to full scene images in order to further validate the research presented here. The reason for carrying out this validation, as stated in the opening chapters of this thesis, is that for an automatic image registration system to be useful to an end user, the algorithms must produce reliable results for full scene images as well as small images.

Many previous studies (Twu, 1996; Boardman et al., 1996; Heipke, 1996) and a lot of modern automatic geocorrection software (PHODIS, SOCET SET) have used image pyramids in the registration of large images. In pyramid matching the full scene images are reduced in resolution by a number of different scale factors to create the layers in the pyramid. Initially the lowest resolution images (which represent the top layer of the pyramid) are approximately aligned using feature matching. The registration results are then passed onto the next layer in the pyramid to improve the alignment of the images. The matching and registration are repeated, and each time the registration results are passed on to the next layer. This process is repeated, and the registration results become progressively more accurate until the bottom layer of the pyramid is reached and the images are accurately registered.

This approach to full scene image registration was tested with SAR and SPOT data, but it was found that it did not produce consistent results. The problem was associated with the feature extraction and feature matching steps in the procedure. It has been stated on numerous occasions in this thesis that the greatest stumbling block to feature based image registration for SAR and SPOT images is the extraction of a suitable number of

corresponding features for the matching to be considered reliable. The use of image pyramids merely magnifies the problem since the extraction process has to be repeated many more times. Therefore, until feature extraction algorithms have been developed which can consistently and reliably extract a large number of corresponding features from the images being matched, pyramid matching cannot be considered a suitable method for registering multisensor data sets.

Since pyramid matching cannot be used in this case, this study proposes to use a method based on image tiles. The full scene images are approximately aligned with each other, and then split into a number of small tiles. Matching then takes place on a tile by tile basis: features are extracted and matched for each individual tile. When all the tiles have been processed the tie points can be combined into one group and used to generate the parameters of the transformation function.

The method of automatic registration of full scene images proposed in this study is a three step procedure:

3. initial approximate alignment using ephemeris data or manually selected tie points;
4. accurate registration using matched patches; and,
5. refined registration using matched edges.

In the first step the full scene images are approximately aligned. This can be done by either using the ephemeris data supplied with the images, or with four manually selected tie points. Results of both methods are illustrated in the next section. After this initial registration, the images are tiled, and each pair of corresponding tiles are registered using the patch matching procedures described in chapter 6. The quantity and distribution of the tie points can then be increased by using the edge matching refinement technique described in chapter 7. Finally, the tie points generated from the tiles are used to register the full scene images.

The images chosen to test this method of full scene image registration are those from which the small test scenes were taken. They depict an area of the southern coast of France close to where the River Rhone flows into the Mediterranean.

8.2 Pre-registration processing

8.2.1 Introduction

Pre-registration processing includes the initial alignment of the full scene images, the cropping of the images to ensure that there is 100% overlap of the images, and the creation of the image tiles. Two techniques are compared for the initial alignment: one using ephemeris data and the other using tie points. The first method is automatic and requires no user intervention, but the second is obviously manual. However, the other pre-registration processes (cropping and tiling) are both automatic.

8.2.2 Initial alignment

With both SAR and SPOT images information is provided describing the locations of the four corners of the images in ground co-ordinates. If these points are used as tie points, then each image can be registered to a ground co-ordinate system using an affine transformation. So long as both images are in the same ground co-ordinate system they can be overlaid on top of each other and considered to be registered to each other. However, this ephemeris information is only provided as a guide to help users visualize the coverage of the image on the ground. Therefore, it is not surprisingly that approximate alignment using ephemeris data produces quite poor results which are worse than those produced using manually selected tie points. Figure 8.1 shows the SAR scene registered to an orthographic ground co-ordinate system using ephemeris data. Overlaid on top of the image are the outlines of the main features taken from the SPOT scene. These features have also been registered to the same ground co-ordinate system.



**Figure 8.1 Approximate registration using ephemeris data
(orthographic ground co-ordinate system)**

It can be seen that over most of the image the alignment is quite poor, although in a couple of places it is not too bad⁶. The misalignment is obviously due to small inaccuracies in each of the rotation, scale factor and translation components of the affine transformation function.

The alignment generated using four manually selected tie points is much more accurate. Figure 8.2 shows the result of registering the images using four tie points placed close to each of the corners of the SPOT image. This time the result is displayed in SAR image co-ordinates rather than ground co-ordinates.

⁶ Even a stopped clock is correct twice a day!



Figure 8.2 Approximate registration using four manually selected tie points (SAR image co-ordinate system)

It is clear that although the images are aligned much more closely, that alignment is still not perfect, which proves that more refined matching at full resolution is very important. Even so, the outlines of the main features from the SPOT image can be seen to approximately align with the features in the SAR image. This alignment should therefore provide a good starting point for the automatic image registration.

A final comment should be made about how this lack of automation affects the validity of this study. Clearly it would be better if this initial step could be fully automated, but it was not considered essential since firstly, the process of locating four tie points is quite straightforward for any end user, and secondly this step gives the end user the opportunity to examine the images before any processing begins. This would be useful for highlighting blunders, such as the incorrect choice of images which do not overlap. Furthermore, if necessary, additional processing could be carried out on the automatically aligned images to try to improve the accuracy of the registration, but in the context of this study, this was seen as unnecessary.

8.2.3 Creation of image tiles

Once the images have been aligned they can be easily cropped so that both images cover the same area on the ground, and are made up of the same number of pixels. Creation of image tiles is also a very simple procedure: the image can either be split into a fixed number of tiles which are all the same size, or a tile size can be chosen, and as many as possible fitted into the image. In this case the latter method was used: tiles of 512×512 pixels were fitted into the full scenes. This resulted in the creation of 99 tiles (9 rows \times 11 columns), one row of which were slightly smaller than the rest of the tiles. The results are shown in figure 8.3 for the SAR image and 8.4 for the SPOT image.

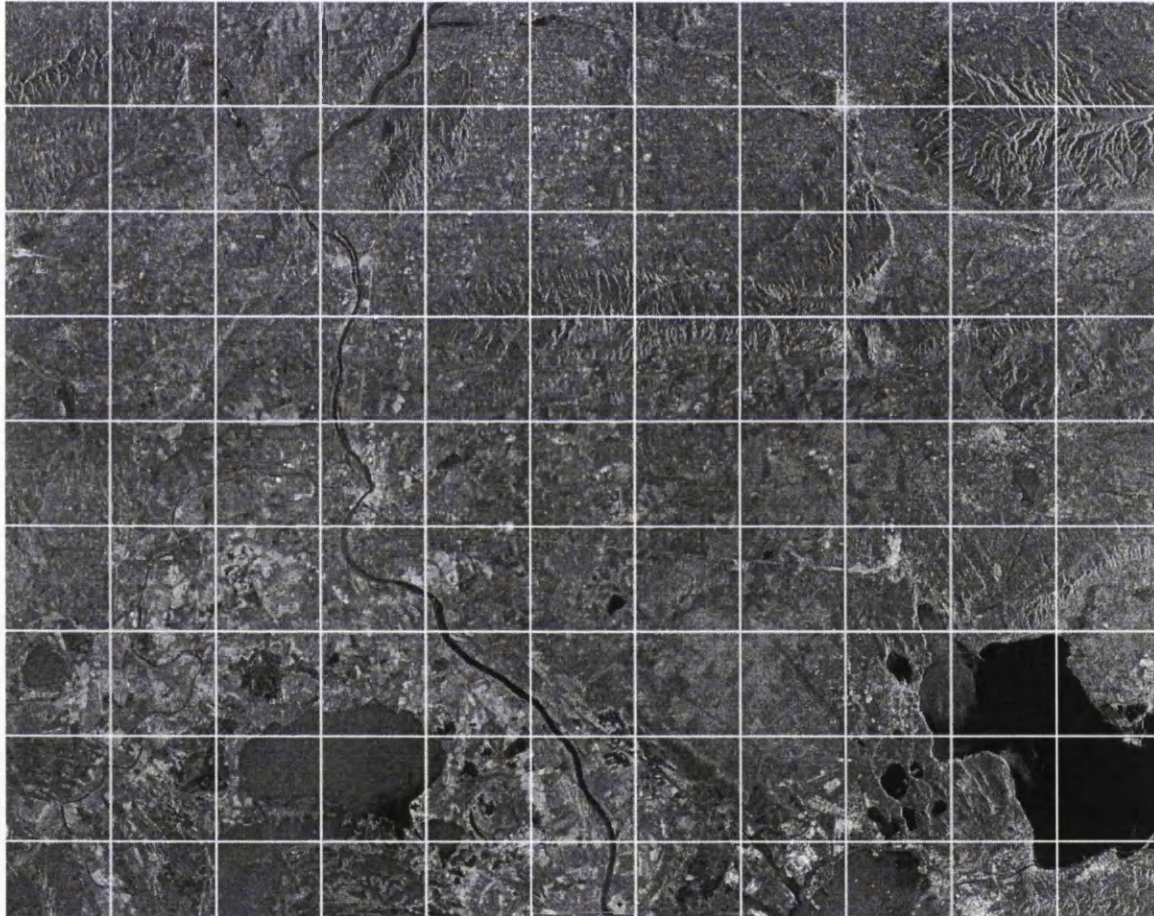


Figure 8.3 SAR image split into tiles

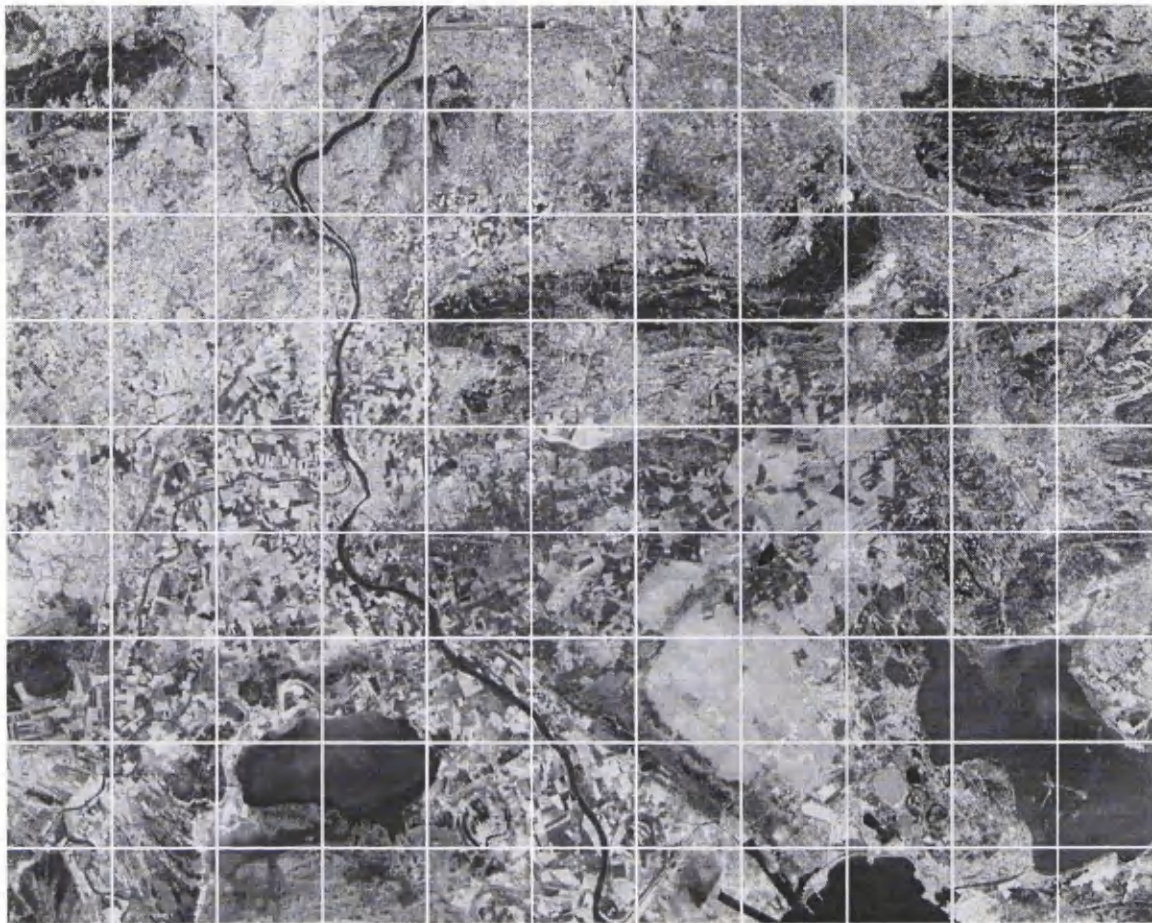


Figure 8.4 SPOT image split into tiles

8.2.4 Conclusions

There is very little which can be said in conclusion to the work presented in this section, except that it is clear that manual approximate alignment of full scene images is more accurate than automatic alignment which relies on ephemeris data. Future research could investigate this issue, but for this piece of work it was deemed unnecessary. The processing of the full scene images described here means that they are now ready to be registered automatically using feature extraction and matching.

8.3 Automatic registration using feature matching

8.3.1 Introduction

This section describes the extraction and matching of features from the tiles created from the approximately aligned full scene images, and the subsequent registration of those images. Exactly the same methods were employed in this section as were described in chapters 6 and 7. The brute force technique of parameter selection for patch extraction algorithms was used, which meant that two of the segmentation algorithms (RWSEG for SAR and OPTISEG for SPOT) could not be used in the feature extraction stage. As was explained before, this is because these algorithms produce results which are finely dependent on multiple parameters, meaning automatic parameter selection is

very difficult. The images were registered using the patch matching results, after which the results of the patch matching were refined using the edge matching techniques described in chapter 7. The images were then registered again to see if there was any improvement in accuracy. This section also includes a discussion of the registration results, and a proposal for using the automatically generated tie points in the orthorectification of the SPOT data.

8.3.2 Polygon extraction and matching

From a quick examination of figures 8.3 and 8.4 it appears that there are only a limited number of tiles which will be of any use in the patch matching procedure. These are tiles which contain strong features in both the SAR data and the SPOT data. The majority of the tiles do not contain sufficient features to guarantee successful patch matching. However, it should be remembered that it is not necessary to match features, and therefore acquire tie points, in all of the tiles. Just enough tiles should be used to ensure that the tie points are distributed fairly evenly across the image, with some located near each of the corners.

For this reason it was decided to manually select and process only those tiles which would have the greatest probability of yielding matched patches. Obviously, it is not essential to include this manual step in the processing chain, since the feature extraction and matching algorithms could have been applied to all the tiles. But by intervening in the automatic processing chain at this point meant an enormous saving in processing time. Future research into the development of this type of automatic registration system could investigate the possibilities of automating this step, but that is beyond the scope of this thesis.

The images were examined and 12 tiles were selected to be used in the registration process. The aim was that these tiles should be spread very evenly across the image, but this was not easy since the principle features which were likely to give good matches were not spread evenly across the images. The selected tiles are shown in figure 8.5 below.

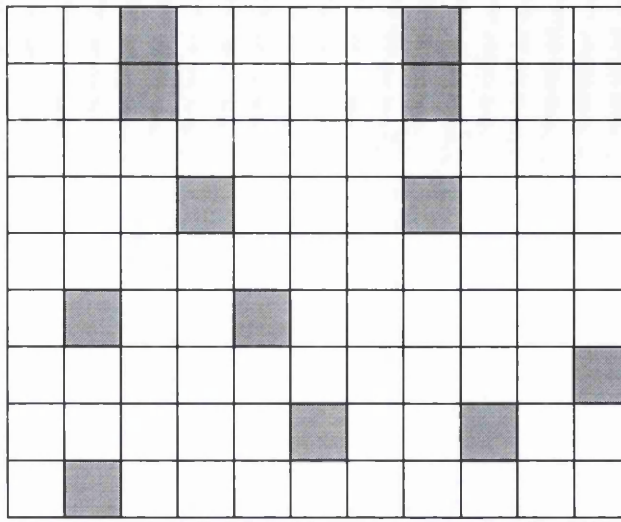


Figure 8.5 Distribution of tiles selected to be matched

The 12 tiles were selected so that some occupied areas as close to the corners of the images as possible, whilst others were spread relatively evenly across the central region of the images. The distribution shown in figure 8.5 indicates that these criteria were almost met – apart from a gap in the top right of the images, the distribution is fairly even.

For each pair of corresponding tiles, patches were extracted and matched using the methods described in chapter 6. The matched patches were extracted using a range of feature extraction algorithms. The details of the extraction and matching processes have not been included here, but it should be noted that all the improvements to the patch matching procedure (described in § 6.4) had to be implemented in order to achieve the maximum possible number of matches. Even so, of the 12 selected tiles, only ten produced correctly matched patches – two of them failed to locate any common features. Figure 8.6 shows the number of correct matches found for each tile that was processed.

		5					3			
		4					2			
			3				0			
	5			2						
										3
					6			6		
	0									

Figure 8.6 Number of correct matches found for each processed tile

As a result of extracting and matching patches from the 12 tiles a total of 39 matched patches were located. However, since two of the tiles failed to produce any matches, the distribution of those tie points is now much worse than the distribution of the original 12 patches. There are no matches in the bottom left corner of the images, nor are there any matches in the region to the right of the centre of the images. It is likely that this distribution of tie points will adversely affect the registration results, but it is a problem which is very difficult to solve. A good distribution of tie points is completely dependent upon the ability to extract and match features across the whole image. With the feature extraction algorithms introduced in this study it has clearly not been possible to do this, partly due to the quality of those feature extraction algorithms, and partly due to the lack of common features in the SAR and SPOT images. This problem is fundamental to automatic feature based image registration, and was introduced in § 2.4.6. It is discussed further in chapter 9.

Figures 8.7 and 8.8 show the matched features for the SAR and SPOT images respectively. Although there are problems with the spatial distribution of the tie points generated from these matched patches, it is still possible to proceed with the registration of the full scene SPOT image to the full scene SAR image. The details and results of this registration process is discussed in the next section.

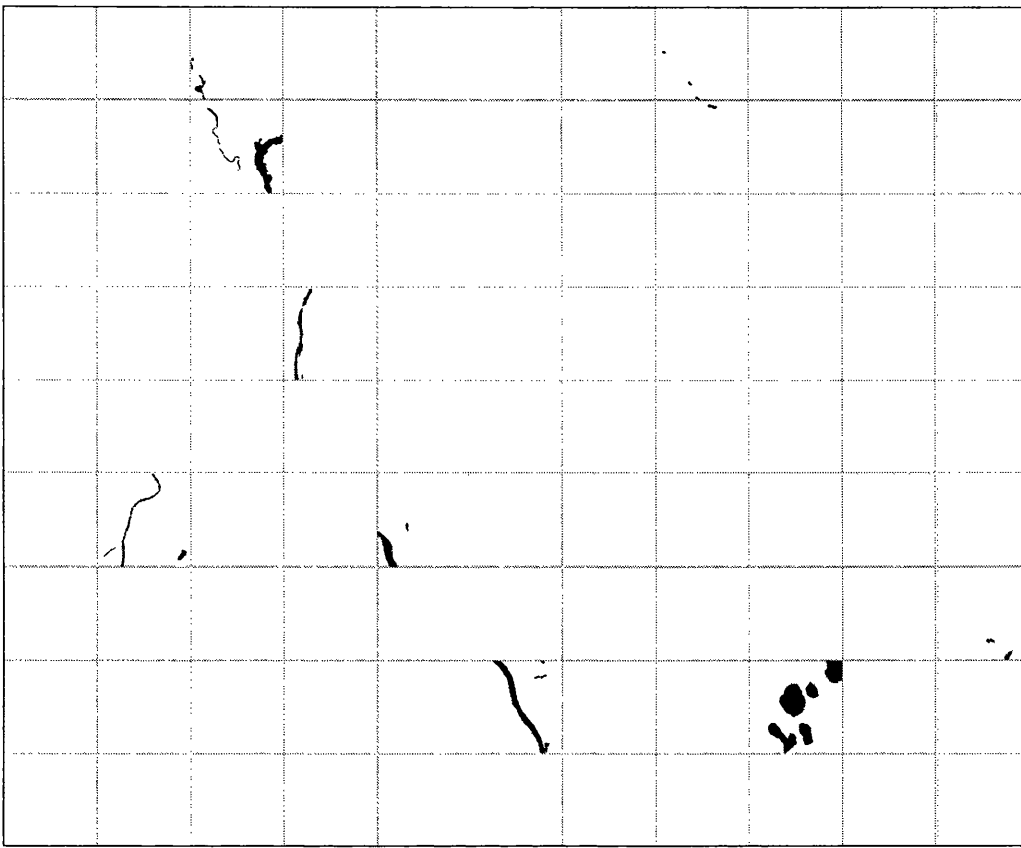


Figure 8.7 Matched features from SAR image

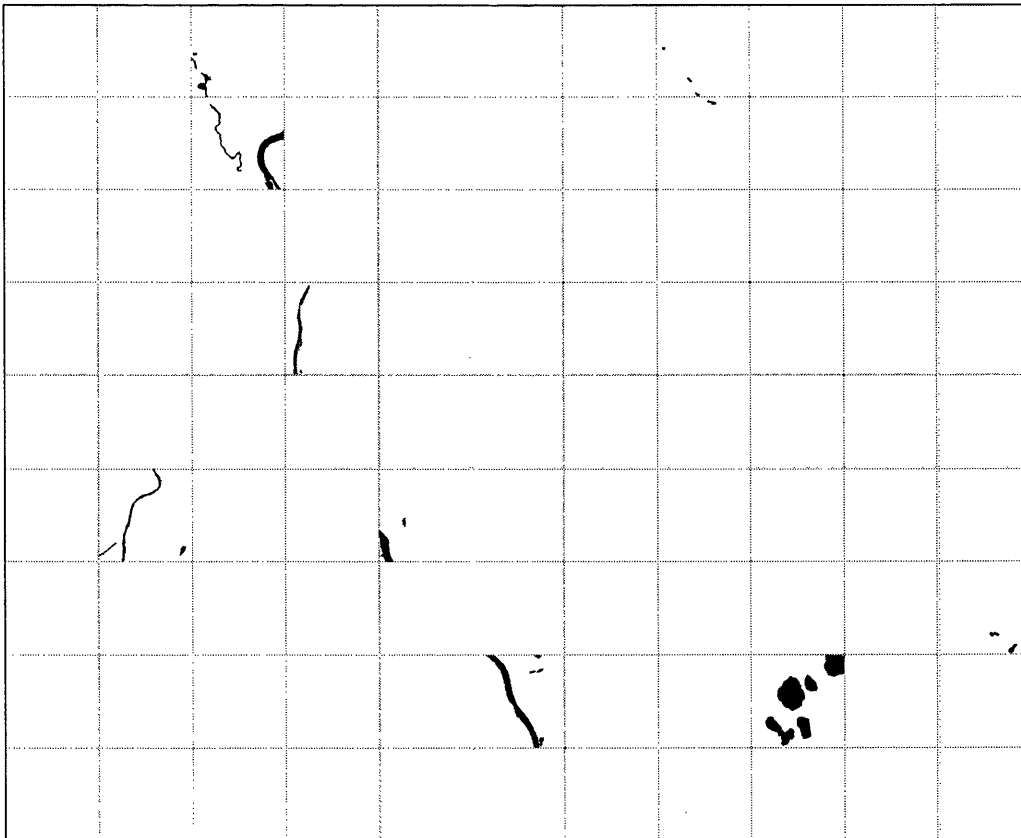


Figure 8.8 Matched features from SPOT image

8.3.3 Image registration

The centroids of the matched patches were grouped together and used as tie points to determine the parameters of a single affine transformation function to register the SPOT image to the SAR image. The parameters of the transformation function were generated using a least squares method. To illustrate the results visually, figure 8.9 shows some patches extracted from the SPOT image overlaid onto the SAR image using this transformation. The quality of the result can be assessed by comparing this figure with figure 8.2 above.

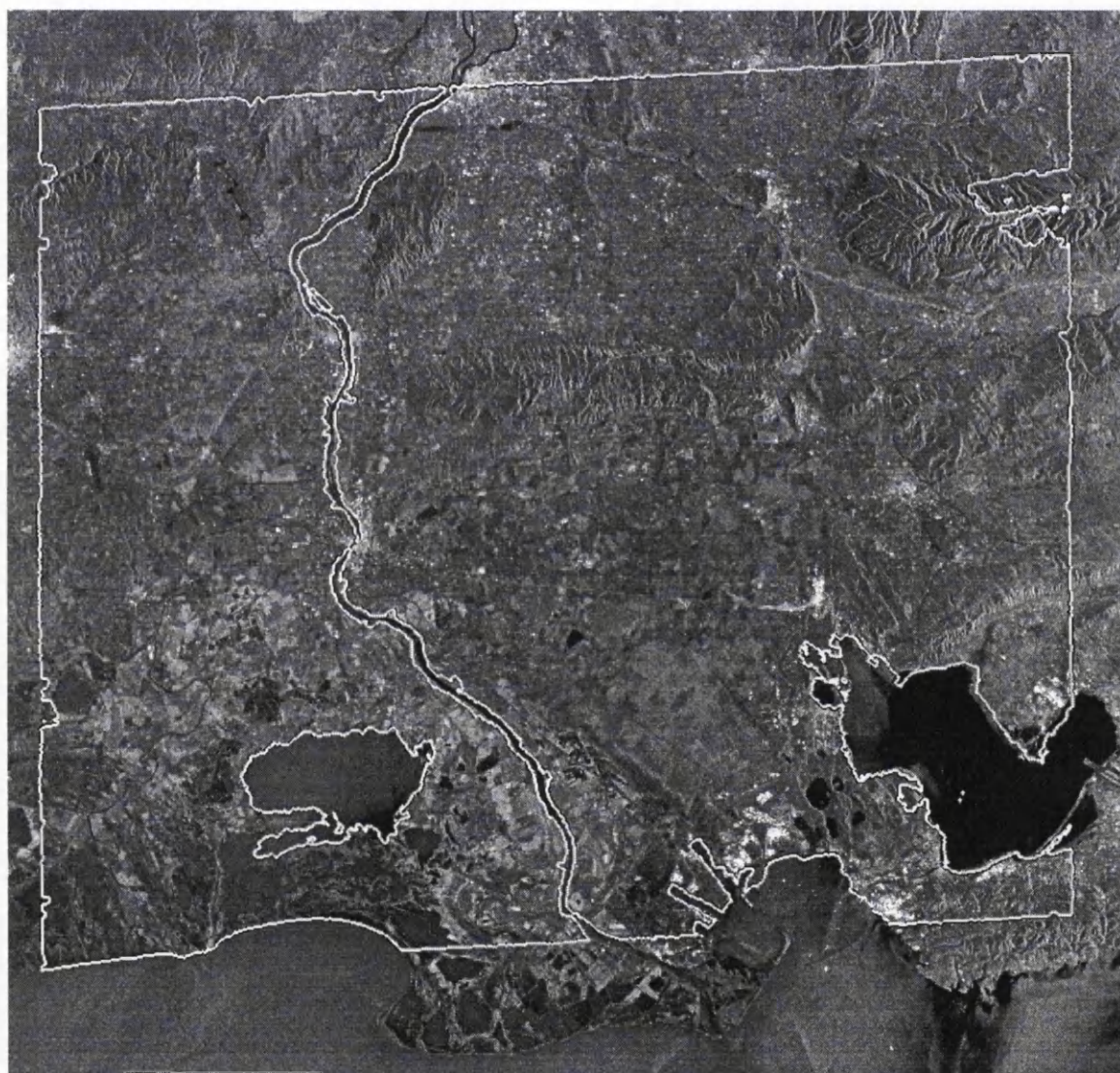


Figure 8.9 Visual result of automatic registration using patch matching

By comparing figures 8.2 and 8.9 it is clear that the patch matching registration technique improves the registration result in some areas of the image, but not in others. It can be seen that there is now better alignment of the whole of the river feature, whereas before with the manual registration there is a marked misalignment at the southern end of the feature. However, the lake on the right of the image does not align as well as before. The most likely cause of this misalignment is that an affine

transformation function has been used to correct for non-linear errors, especially those caused by terrain, which it is not able to do. Therefore, it would be unreasonable to expect that the registration result could be much better than it actually is.

Obviously, it is insufficient to judge the results by interactively observing the reduced resolution images. In order to statistically assess the accuracy of the registration, the same procedure was employed here as was used in chapter 7. The tie points located by patch matching were split into two groups: one group was used to generate the parameters of the transformation function, and the other group were used as check points.

Test 1		Test 2	
<i>x</i> residual (pixels)	<i>y</i> residual (pixels)	<i>x</i> residual (pixels)	<i>y</i> residual (pixels)
-9.40	-10.08	6.12	0.06
-6.98	-4.79	-9.08	-12.15
9.72	1.65	-0.97	-0.68
-4.54	21.04	16.68	-0.07
-14.21	5.51	-4.96	4.15
-10.28	2.87	-6.28	1.81
5.77	-17.71	16.25	-12.52
-10.38	3.74	7.57	3.83
-14.38	8.73	-10.03	3.29
-12.57	6.62	-6.29	-4.62
6.53	-7.87	12.61	0.59
10.40	0.31	15.92	0.76
-31.20	-0.06	17.11	0.19
-33.48	-0.96	-24.50	1.31
13.33	1.92	18.25	-3.07
15.41	0.11	13.22	1.98
16.69	2.40	19.88	-0.22
-2.11	1.52	-7.52	0.02
1.90	-0.75	2.26	1.49
2.15	-0.58	2.73	1.35

Table 8.1 Residuals in *x* and *y* for two sets of check points

The check points were transformed and residuals were calculated by comparing the predicted location of the check points with their actual location. In order to confirm the results, the process was then repeated, but with the two groups interchanged: the check points were now used as tie points, and the tie points were used as check points. The residuals for the two tests are shown in table 8.1 and the RMS residuals are shown in table 8.2. In x , the residuals range from approximately -33 to 20 pixels, and in y from -17 to 21 pixels, but the means of both x and y are close to 0. The y residuals are much smaller than the x residuals. The results for test 2 are very slightly better than those for test 1, but the difference is small and not significant.

	Test 1	Test 2
RMS residual in x	14.17 pixels	12.65 pixels
RMS residual in y	7.52 pixels	4.43 pixels
Total RMS residual	16.04 pixels	13.40 pixels

Table 8.2 RMS residuals for two sets of check points

Table 8.2 shows that the RMS residual is significantly greater in the x direction than in the y direction for both of the tests. An explanation for this is given below. A quick glance at these residuals would imply that the registration result was really not very good at all, but considering that an affine transformation has been used to register to full scene multisensor images, this is hardly surprising. An important aspect of these results is the fact that the residuals for test 1 are similar to those for test 2, implying that both combinations of tie points and check points are representative of the whole set.

Both groups of residuals listed in table 8.1 have been plotted on the original SAR image so that their distribution can be visualized. The result is shown in figure 8.10, but note that the magnitude of the residuals have been increased by a factor of 80 so that they are easily visible. Residuals from test 1 are shown in pink and residuals from test 2 are shown in blue.



Figure 8.10 Two sets of residuals overlaid on original SAR image

Figure 8.10 shows that the distribution of residuals for both groups is not completely random. There is clear systematic error in the x direction, but this is not necessarily the case in the y direction. This pattern is most likely caused by the fact that a linear transformation function has been used to correct for non-linear distortions in the image. The affine transformation can correct for differences in rotation, translation and scale, but the images being registered incorporate more complicated distortions, due mainly to perspective and terrain effects. By comparing the images with a topographic map of the region, it was estimated that the tie points had a vertical range of about 50m. For an object with a height of 50m being imaged off-nadir (i.e., with a look angle of 25°) by the SPOT satellite, the radial displacement of the top of the object from its true orthographic location is approximately 27m, or three pixels. For an ERS SAR image, the corresponding displacement is approximately 46m, or four pixels. Therefore, the maximum error due to terrain effects could be as much as seven pixels. These distortions will be much greater in the x (across track) direction than in the y (along track) direction, due to the way in which the images are acquired. The residuals displayed in figure 8.10 confirm that the RMS errors in the x direction are greater than the RMS errors in the y direction. Therefore, although the residuals may seem comparatively large, they do actually reflect the physical situation.

It would be expected that the residuals generated by the affine transformation are quite high since a first order transformation function cannot model the distortions that exist between full scene SAR and SPOT images. However, just because the residuals are high does not mean that the tie points which have been located in the images are inaccurate. The accuracy of the tie points can be verified by observing the matched patches in each tile. Figures 8.11 to 8.20 show the matched patches from each tile overlaid on top of the original SAR and SPOT images. In each case it can clearly be seen that all the patches have been matched correctly. Tiles are numbered according to their row/column location in the full scene images.

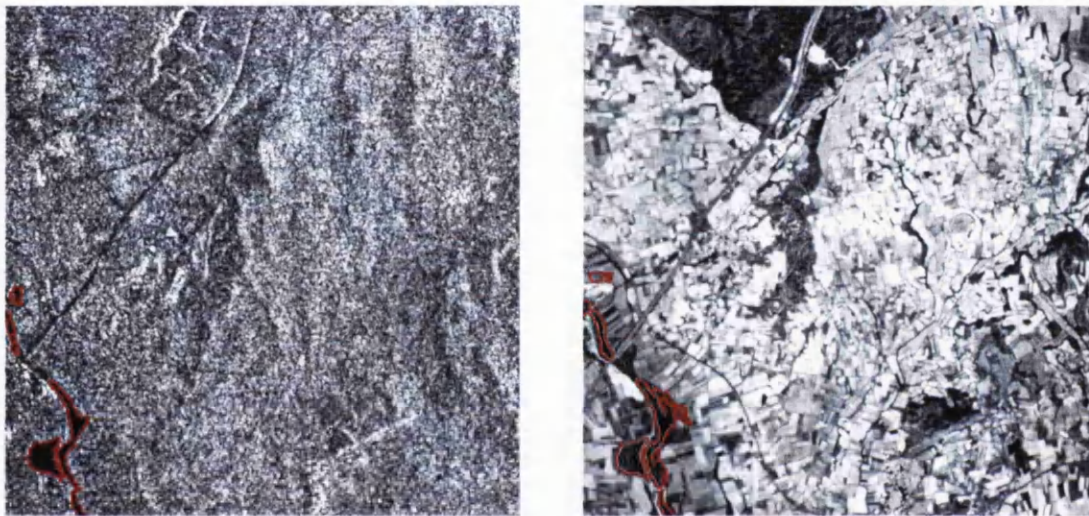


Figure 8.11 Matched patches from tile at (1,3)

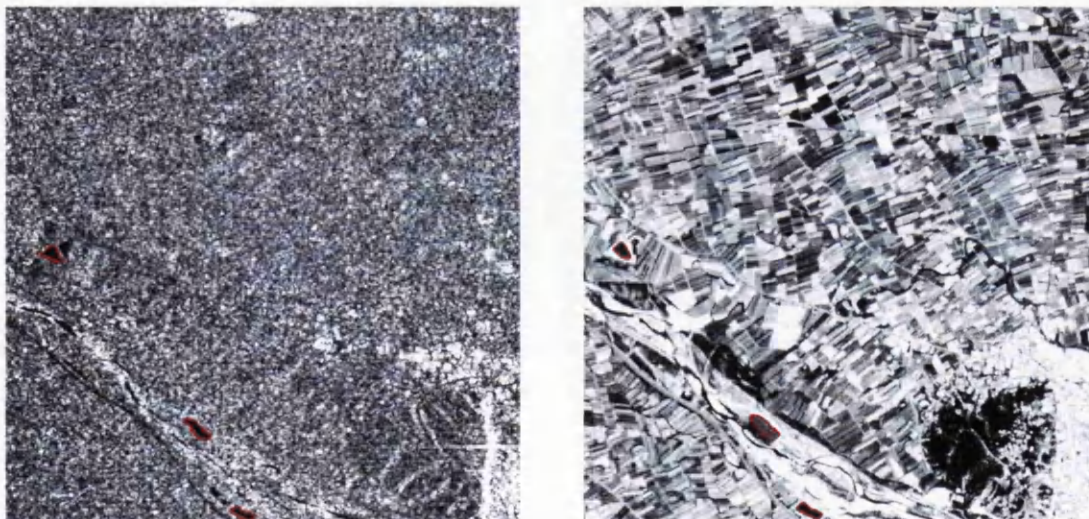


Figure 8.12 Matched patches from tile at (1,8)

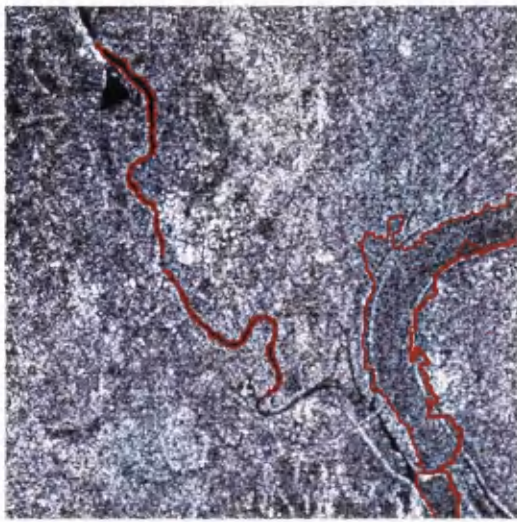


Figure 8.13 Matched patches from tile at (2,3)

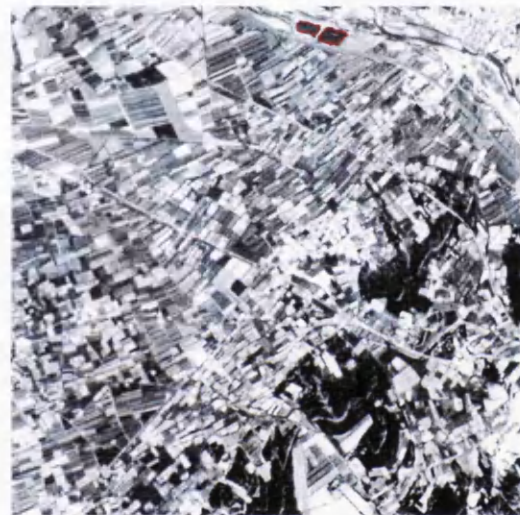


Figure 8.14 Matched patches from tile at (2,8)

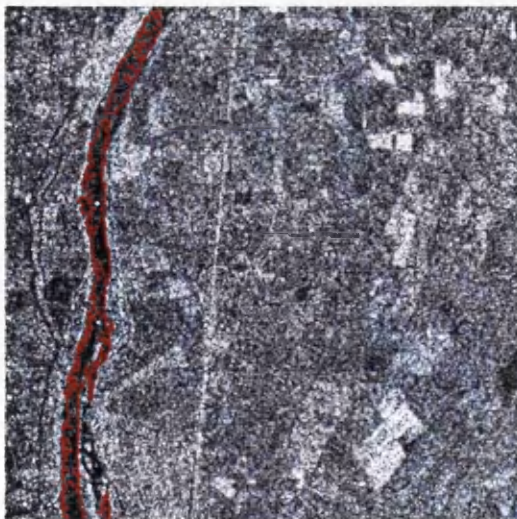


Figure 8.15 Matched patches from tile at (4,4)



Figure 8.16 Matched patches from tile at (6,2)

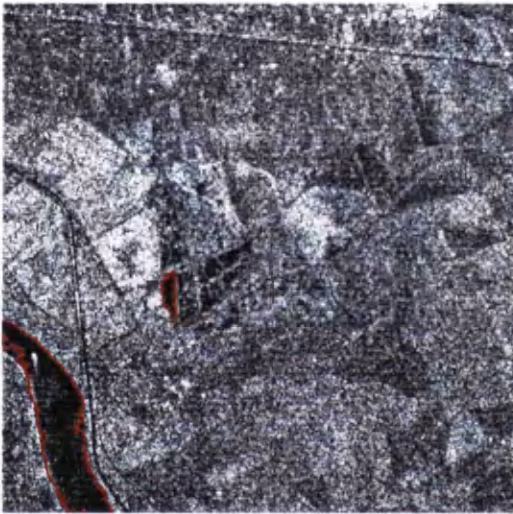


Figure 8.17 Matched patches from tile at (6,5)

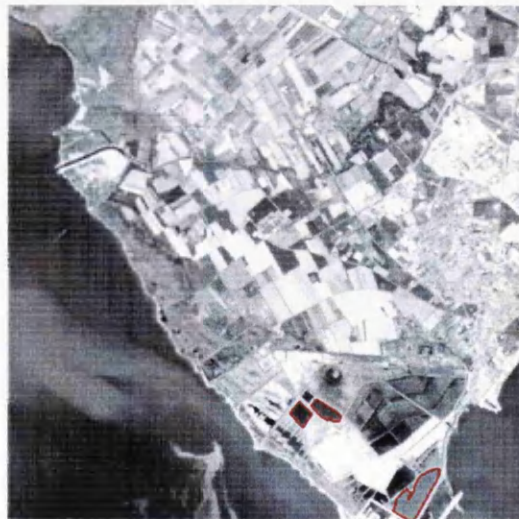
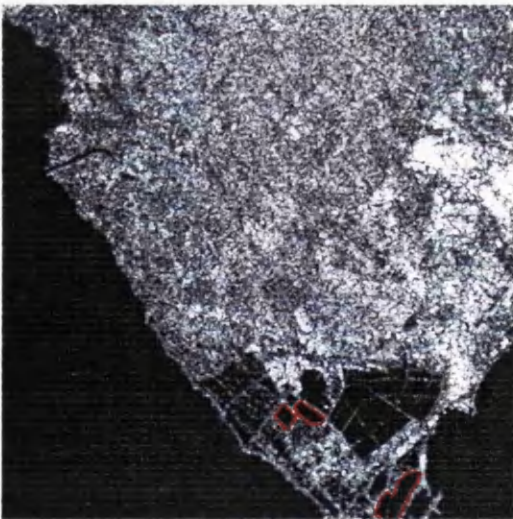


Figure 8.18 Matched patches from tile at (7,11)

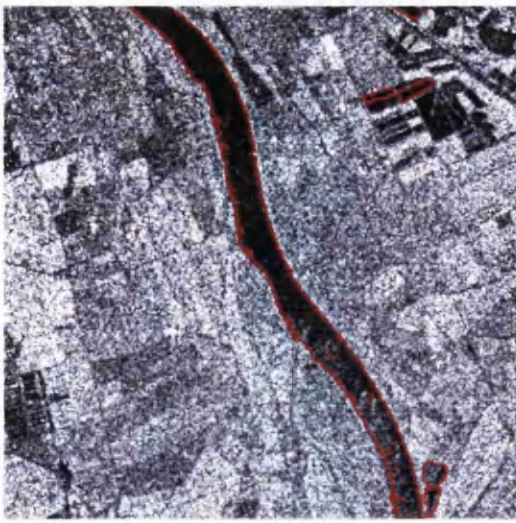


Figure 8.19 Matched patches from tile at (8,6)

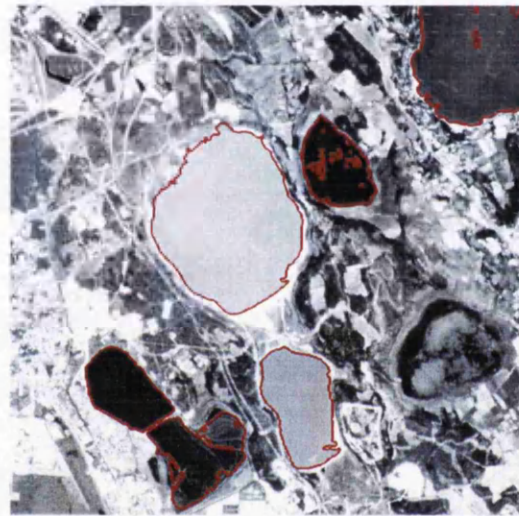
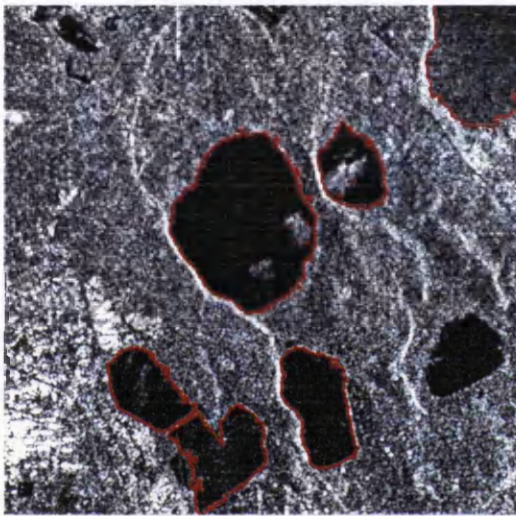


Figure 8.20 Matched patches from tile at (8,9)

Thus it can be seen from the ten pairs of images above that all the patches have been correctly matched, and therefore all the tie points can be considered to be reliable. However, the images also illustrate how errors can occur in the exact location of some of the tie points. The errors arise because the centroids of the matched patches are used as the tie points, but these centroids are not always reliable. In cases where the matched feature is in contact with the edge of the tile, the misalignment between the two tiles can cause the feature to be a slightly different size and shape in the two images. Consider the river feature in figure 8.17 for example: due to the misalignment of the two images, the feature is slightly bigger in the SAR image than in the SPOT image. The consequence is that the location of the centroid (in ground co-ordinates) in the SAR image will be different to the location of the centroid (in ground co-ordinates) in the SPOT image, resulting in a less accurate tie point. These less accurate tie points contribute to RMS residuals detailed in table 8.2. In order to assess the accuracy of the tie points it would be necessary to compare the location each one in ground co-ordinates

and determine the discrepancies at each point. However, since no ground control data is available it is not possible to do this.

This problem of inaccurate tie points could be solved by manually eliminating dubiously matched patches, but this would require a lot of user intensive processing and would leave even fewer tie points with which to register the images. A much better solution is to refine the quality of the tie points using edge matching, which is described in the next section.

Another reason for the comparatively high residuals is the relatively uneven spatial distribution of the tie points. Previous chapters have stated the importance of an even distribution of tie points, but this has not been achieved here due to the lack of matched features. Therefore this will also contribute to the magnitude of the residuals.

8.3.4 Refinement using edge matching

Although the patch matching results have already been used to register the full scene images, it is still useful to investigate the possibilities of refining the results using the edge matching techniques described in chapter 7. It is unlikely that the RMS residuals shown in table 8.3 can be significantly improved, since an affine transformation will still be used for the registration, but a small improvement is expected since the edge matching should eliminate errors caused by poorly matched centroids. If a larger set of more accurate tie points can be found using edge matching, then this may prove useful for the method of automatic registration proposed in the next section.

In order to refine the results using edge matching, each tile which includes three or more matched patches must be registered. The important point to note is that although the affine transformation will be used for each tile, the tie points will be different, and therefore the parameters of the transformation will be different. By working with individual tiles, rather than all the tiles collectively, the registration results for each tile will be more accurate since the transformation will only be correcting for local distortions. For each tile the method proposed in chapter 7 was followed. Both SAR and SPOT tiles were smoothed using the median smoothing filter, and edges were located using the Sobel operator. Unwanted edges were masked out, leaving only those edges near the boundaries of the matched patches. Edges were matched, tiles were registered, and the corresponding results are shown in table 8.4. There is no need to register individual tiles since only the tie points are needed, but RMS residuals do give a good indication of the accuracy of the matching of the edges.

Tile (row, column)	No. of tie points found	RMS residual (pixels)
(1, 3)	28	3.02
(1, 8)	84	2.60
(2, 3)	955	1.78
(4, 4)	271	2.27
(6, 2)	423	1.81
(7, 11)	50	12.37
(8, 6)	752	10.40
(8, 9)	925	1.86

Table 8.4 Number of tie points found and RMS residual for each processed tile

Table 8.4 illustrates the range of results which can be obtained using edge matching, depending on the content of the tiles. Tiles with a large number of strong features will yield a greater number of matched patches, and more matched patches means more edge pixels which can be matched. Thus the number of tie points found for each tile, shown in table 8.4, can vary greatly. In addition the accuracy of the registration can be found to vary a lot as well. This is due to a number of factors such as terrain effects, perspective distortions and distributions of tie points, which will be different for all tiles. Most of the tiles yield RMS residuals of 3 pixels or less, but two of the tiles produce much poorer results. This does not necessarily imply that the tie points are incorrect (although this could be a factor), but distortions from other sources may be more prominent in these tiles. Therefore, although the RMS residuals for each tile can be used to highlight potential sources of error in the full scene registration, they should not be used to remove tie points from the full scene registration.

The result of performing the edge matching is that for each tile a much larger set of tie points has been found for each image. However, although the SAR component of these tie points are in the same co-ordinate system as all the other SAR components, the SPOT components of the tie points are in different co-ordinate systems. This is because each SPOT tile has been registered to its corresponding SAR tile slightly differently. Therefore, in order to be able to use these tie points to register the full scene images, the SPOT components of the tie points have to be transformed back into their original co-ordinate system. Once this is done (simply by reversing the transformation function derived by patch matching) all the new tie points can be combined into a single group and used to register the full scene SPOT image to the full scene SAR image.

A total of 3488 tie points were found using the edge matching technique. Their spatial distribution across the image was similar to the tie points found using the patch matching technique (i.e. sufficient but not perfect) but obviously this time the points were much more clustered. This distribution is shown in figure 8.21.

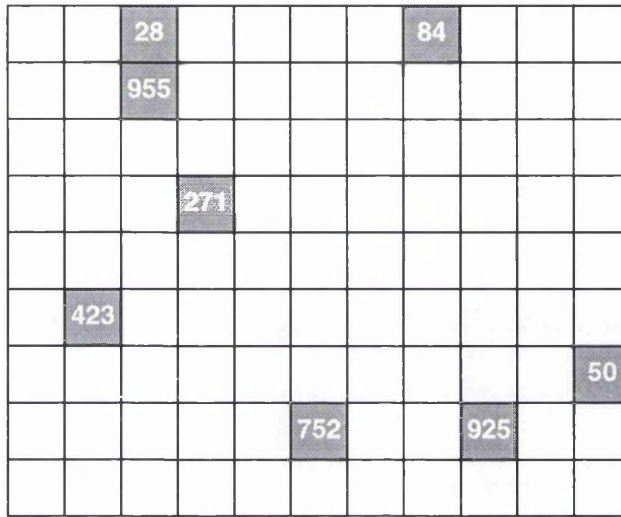


Figure 8.21 Spatial distribution of tie points found using edge matching

These tie points were used to register the images. As with previous tests, the tie points were split into two groups: one for calculating the registration parameters and the other to act as check points. The RMS residuals are shown in table 8.5.

	Test 1	Test 2
RMS residual in x	9.89 pixels	9.84 pixels
RMS residual in y	4.47 pixels	4.50 pixels
Total RMS residual	10.85 pixels	10.82 pixels

Table 8.5 RMS residuals calculated using tie points from edge matching

As was anticipated, the refined registration parameters determined using edge matching has produced residuals which are superior to those determined using patch matching alone, but not significantly different. Since there are many more tie points now (3844 compared to 39), the results for the two tests (test 1 and test 2) are much more similar. It is not necessarily true that the residuals are smaller since more tie points have been used to calculate the transformation parameters – table 8.4 shows that in this example there is no correlation between number of tie points and the magnitude of the RMS residual. The residuals are similar to those determined using patch matching since the patch matching procedure has played a role in locating the tie points used for edge matching. The x residual is larger than the y residual (by approximately 5 pixels), but as with the

patch matching results, the terrain and perspective effects can account for the difference. It is interesting to note that the total RMS residual of approximately 11 pixels is greater than most of the RMS residuals for each individual tile. Two reasons can be used to explain this fact. Firstly, the spatial distribution of residuals across the tiles is in general much more even (and therefore better) than the spatial distribution of the tie points across the whole full scene image. Secondly, distortions due to terrain and perspective effects are local to the tiles, and can therefore be more accurately modelled with the affine transformation. If it is assumed that errors due to poorly matched centroids have been removed in the refinement process, and that the increased number of tie points has no effect on the overall error, then the lower RMS residual that has been achieved must be due to terrain and perspective effects, and the clustering of the tie points. As before, the residuals were plotted on the SAR image in order to visualize their distribution (see figure 8.22).

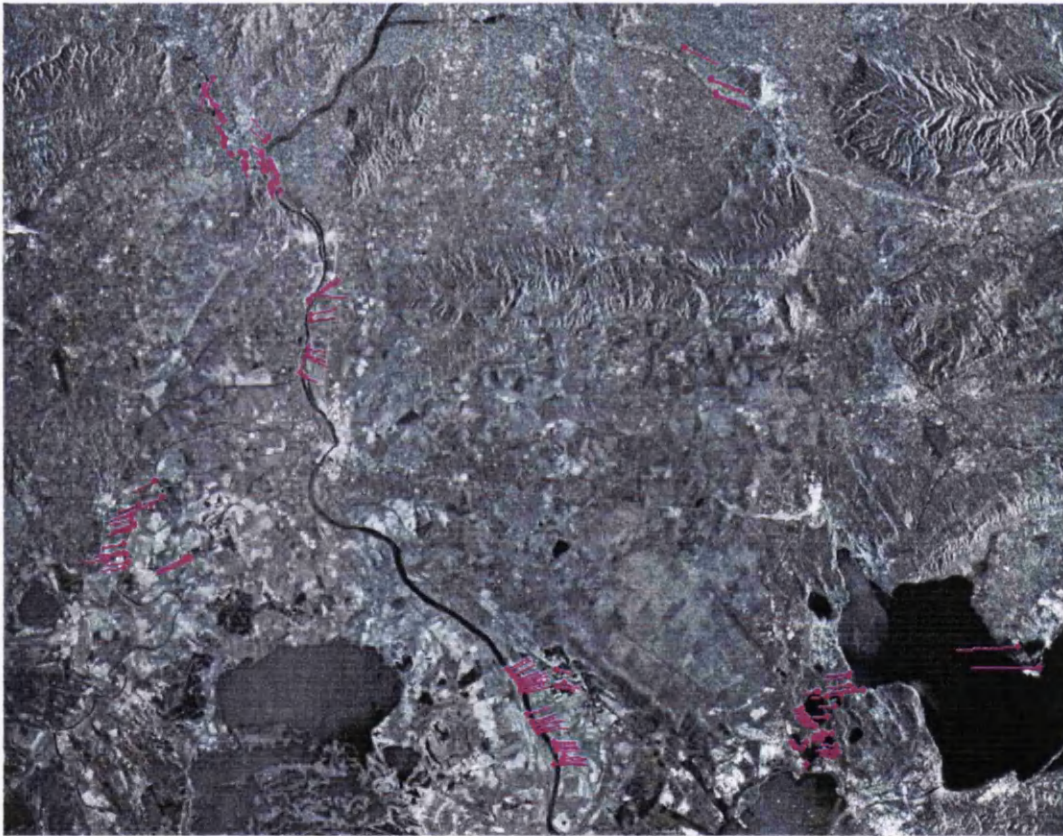


Figure 8.22 Refined residuals from edge matching overlaid onto SAR image

The magnitude of the residuals in figure 8.22 have been enlarged by a factor of 80 for clarity. Also note that only the results for one of the tests have been plotted (unlike before when the results of the two tests were plotted), and that for this test only 5% of the residuals have been plotted.

It can be seen that the spatial distribution of the residuals of the refined registration is very similar to the spatial distribution of the residuals of the original registration⁷, except that there are many more residuals clustered into fewer regions across the image. The distribution of residual orientations is also quite similar. The systematic error is still quite clear since the refinement technique is not able to correct this. Comparison of the magnitude and direction of the residuals with a topographic map does not indicate any obvious correlation with terrain height, but this is most likely masked by the strong clustering of the tie points. A much more even distribution of tie points is essential if residuals are to be correlated with terrain heights.

Since the refined registration results are fairly similar to those achieved using patch matching alone, the overlay of the SPOT patches on the SAR image is not significantly different from the result shown in figure 8.9. Even so, it is shown in figure 8.23 below.



Figure 8.23 Visual result of automatic registration refined using edge matching

⁷ If this were not the case, then the edge matching could not be considered to be a refinement technique.

Figure 8.23 is not significantly different from figure 8.9 since the refined registration is only slightly different from the original. This small change is not possible to detect visually between the two images. The above figure has been included to prove that no blunders have been made in the refinement process, and that the result is still reasonable.

As with the original patch matching results, the relatively large residuals obtained with the refined registration procedure do not mean that the tie points are incorrect – however, it is clear that the transformation function is incorrect. In order to circumvent the problem of selecting a suitable polynomial transformation function, a method for improving the registration results based on the technique of patch and edge matching is proposed in § 8.4.

8.3.5 Conclusions

The automatic registration of tiles making up two full scene images has been presented in this section. The work is very similar to that of previous chapters where small scenes, which were very similar to the tiles used in this chapter, were registered. The same techniques were used, namely patch extraction using brute force parameter selection, patching matching based on patch attributes, and refinement using edge matching.

Features were not extracted from all the patches since redundancy has played a role in the image registration system yet again. This time the redundancy of tiles means that not all tiles have to be matched, therefore greatly reducing the processing time but without compromising the accuracy of the results.

At the patch matching stage, all the tiles yielded too few patches to allow a reliable statistical analysis of the subsequent registration of individual tiles. However, when all the tie points were combined and the full scene images registered to each other, it was possible to calculate some RMS residuals. Analysis of these residuals highlighted the errors caused by using the incorrect transformation function.

The registration results were improved using the refinement method based on edge matching. From the refined registration results it was still clear that the incorrect transformation function had been used, but since more tie points were used in the calculation of the transformation parameters, the overall residuals were better, indicating that the refinement technique is a valuable improvement the registration procedure.

A possible solution to the problem of inaccurate registration caused by the incorrect choice of polynomial transformation function would be to use a piecewise polynomial. If matches could be found in every single tile across the full scene images, this could be a good starting point for determining a piecewise polynomial transformation function which could register the images. Although it would not be able to eliminate the errors completely, it would be a significant improvement on the affine transformation used here.

However, to ensure accurate registration, the tie points found using patch and edge matching could be combined with photogrammetric techniques to perform a full orthorectification of the SAR and SPOT images. This is discussed in the next section.

8.4 Automatic orthorectification of SPOT data

This chapter has illustrated how two full scene, multisensor images can be automatically registered quite successfully using feature matching. However, the quality of the final result is compromised by the simplicity of the transformation function. It seems unfortunate therefore that a lot of effort has been put into automatically locating the tie points, but the accuracy of these tie points is not reflected in the final result. It is obvious that the affine transformation should not be used to register the images, so some more complicated mathematical model should be considered. Using higher order polynomials would give better residuals, but the accuracy of the registration would probably be much worse, so they can be ruled out. However, if a digital elevation model of the whole coverage of the two images were available, and if it was of sufficient resolution, it would be possible to automatically register the images to a high degree of accuracy using well-established photogrammetric techniques. The proposed method for is summarized in the flow chart in figure 8.24.

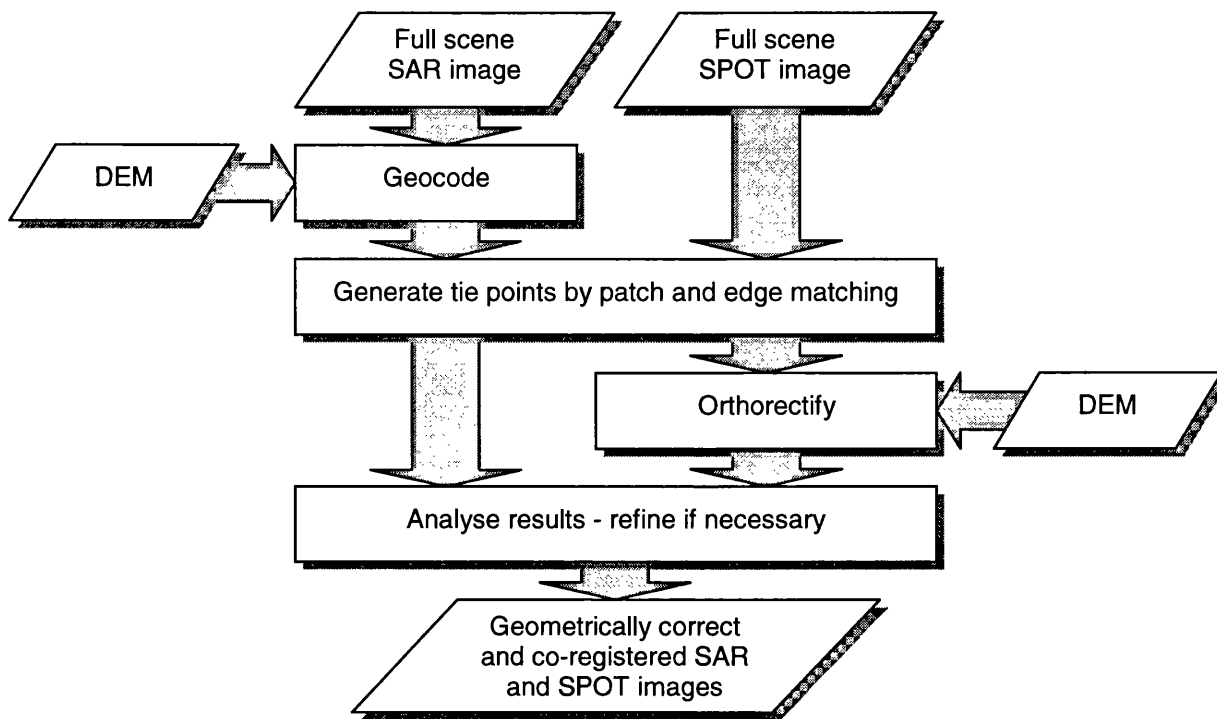


Figure 8.24 Method of automatic registration using photogrammetric techniques

The first step in the procedure is to geocode the SAR data using the DEM. For ERS images supplied with detailed ephemeris data, this can be performed fully automatically with no need for human intervention whatsoever. Other SAR images (such as those from Radarsat) benefit from the use of one or two GCPs. After geocoding, each pixel in the original SAR image can be expressed in three dimensional ground co-ordinates (x , y , z). The second step is to generate tie points between the SAR and SPOT images using the patch matching and edge matching techniques described in this thesis. The result is that a number of points are found in the SPOT image which can be directly associated with points in the SAR image. Therefore, for each of these points located in the SPOT image, the (x , y , z) location is known in ground co-ordinates, effectively making these points ground control points. This means that all the information required to orthorectify the SPOT image is available: a DEM, a set of GCPs and a camera model. Thus, the next stage of the processing is to orthorectify the SPOT image using this data. The result is that both the SAR and SPOT images are now geometrically corrected, and expressed in the same ground co-ordinate system - they are effectively registered to each other. The final stage in the processing is to compare the locations of features in the images in order to assess the accuracy of the registration. Any inaccuracies could be reduced by repeating the patch matching procedure and generating a new set of tie points. These points would then be used to determine an affine transformation function which, when applied to the images, would reduce any small errors that may exist. A similar method of orthorectifying SPOT images using the superior geometry of SAR images was

proposed by Renouard and Perlant (1993). They found that the planimetric accuracy of SPOT products could always be increased when correct with ERS images.

The process described here is automatic in nature, and produces accurately registered full scene images. It does however depend completely on having a sufficiently accurate DEM with a coverage that includes the extent of both of the images. If stereo SPOT images and some GCPs were available, then a DEM could be created automatically using area based matching (Day and Muller, 1989). Alternatively, a DEM could be created from stereo SAR images (Sowter, 1998; Dowman et al., 1997), although this would be less accurate than the SPOT DEM. One further method of generating a DEM would be to use SAR interferometry (Massonet, 1997). Although this method can produce very accurate DEMs, it is dependent on having exactly the right images.

8.5 Discussion

This chapter has shown how the technique of automatic registration based on feature matching, described in the preceding chapters of this thesis, can be applied to full scene images. The procedure follows three distinct steps:

1. initial approximate alignment using four manually selected tie points;
2. accurate registration using matched patches; and,
3. refined registration using matched edges.

An initial approximate alignment technique based on ephemeris data was investigated, but it was found that the alignment was quite poor. The simplicity of selecting four tie points means that pursuing a more complicated but automatic initial approximate alignment strategy seems unnecessary. Therefore, it was proposed that the manual technique should be employed.

The patch matching procedures were successful in yielding a number of matches spread across the images. The quantity and distribution were sufficient to allow the full scene images to be registered and residuals to be calculated, but insufficient to achieve an accurate registration. The magnitude and angular distribution of the residuals could be explained by the spatial distribution of the tie points, and the fact that an affine transformation was used to correct for nonlinear distortions (terrain and perspective effects). An analysis of match patches indicated that although there were no blunders in the matching process, it was likely that some of the tie points may have been inaccurate (to the order of a few pixels). Further refinement of the registration using edge matching

gave improved results by eliminating these inaccurate tie points, but the transformation function and tie point distribution still had a significant effect on the residuals.

The extraction and matching of patches described in this section highlights one of the principle advantages of using tiled full scene images. This is the fact that not all of the tiles have to be matched; this means that the volume of processing can be greatly reduced, but the final registration results should not be adversely affected, as long as a sufficient number of patches can be matched.

The residuals showed that the accuracy of the final registration was dependent on the choice of transformation function. However, it is important to remember that the final registration accuracy will also depend on the spatial distribution of tie points, irrespective of which transformation function is used. Therefore, to ensure accurate registration results, it is essential that patches are matched across the whole area of interest, rather than in just one or two localized regions. Thus the same statement is repeated here as has been made on numerous occasions before in this study: the success of the proposed automatic registration system is directly dependent on the ability to extract and match similar features from both the SAR and SPOT data. The distribution of the patches that were matched in the example described in this chapter was reasonable, but it could have been much better. Matches were located close to the edges of the images, as well as in the centre, but a wider and more even distribution would have been desirable. Further research in feature extraction routines is necessary in order to improve the quantity and spatial distribution of the matches.

A significant difference between registering small images and full scene images is that the refinement technique will only increase the spatial distribution of tie points in the small scenes. The extent to which the spatial distribution is increased is dependent on the size of the matched patches compared to the size of the image. For small images, the patches are comparatively large, but for full scene images, they are comparatively much smaller. Furthermore, since not all tiles could be refined, and only matched edge pixels have been used in the refined registration, the spatial distribution of tie points is actually worse in the refined registration than in the approximate registration determined from matching patches alone.

In answer to the problems associated with the choice of polynomial transformation function, a method of automatic orthorectification of SPOT data using the feature matching results was proposed. In using traditional photogrammetric techniques the

errors caused by polynomial transformation functions can be avoided, but to do so a sufficiently accurate DEM is required.

Chapter 9 DISCUSSION

The opening chapter of this thesis introduced the subject of data integration and explained how it consisted of two procedures: image registration and data fusion. It was shown that data integration is a very important aspect of remote sensing that warrants detailed research. This study set out to investigate the feasibility of automating the image registration component of data integration, when used with multisensor data sets. In order to achieve this, a method of automatic image registration was proposed and subsequently researched. This method, based on patch matching and edge matching, was shown to be successful at automatically registering SAR and SPOT images to a level of accuracy similar to what could be achieved with manual image registration, but in much less time, and with very little human effort. This proposed method of automatic image registration was illustrated with both small (512×512 pixels) and large (full scene) images.

The second chapter reviewed various methods of image registration. As a result of this review it was clear that although the structure of an automatic multisensor image registration system could be well defined, the individual algorithms that make up such a system are not well defined. It was proposed therefore that this study should concentrate on the development of those algorithms. Consequently, five objectives were outlined. To recap, these were:

1. to investigate and develop automatic feature extraction algorithms for use with SAR and SPOT data;
2. to develop an automatic patch matching algorithm for use with features extracted from SAR and SPOT data;
3. to integrate the patch extraction and patch matching procedures;
4. to investigate the possibilities of registration of both small and large images using the patch matching results; and,

5. to investigate the possibilities of extending the matching results to allow for full orthocorrection of the data sets.

Objective 1

Chapter 5 investigated a selection of feature extraction algorithms which could be applied to SAR and SPOT images. The three techniques which were chosen to be used in the patch matching algorithm were automatic thresholding, homogeneous patch extraction and segmentation. The automatic thresholding algorithm was based on the established technique of bimodal histogram splitting, but included a reliable method of locating the optimum threshold level. The homogeneous patch extraction algorithm implemented here was essentially the same as described in previous work. Of the four segmentation algorithms described in chapter 5, two were based on previously published research, and the other two were a part of commercial software packages. Of all six algorithms, only one (automatic thresholding) could operate automatically. The others needed some form of human intervention in order to obtain optimum matching results. Hence it was found that it is not possible to develop automatic feature extraction algorithms in isolation. To achieve automation they need to be combined with the feature matching process.

Objective 2

Chapter 6 proposed and investigated an automatic patch matching algorithm. The algorithm was a three step procedure which used the results of a cost function, the shape of the patches, and the separations of similar patches to find corresponding patches extracted from the SAR and SPOT images. Although the idea of the matching algorithm came from previous research, the procedure described here was essentially new. It was found that matches could be detected as long as similar features were extracted from each image. As a result it became clear that if the patch matching algorithm were to be universally successful, it would have to be combined with the patch extraction stage in the processing. Even so, the algorithm was considered to be a success since all the matches it detected were correct; there were no blunders.

Objective 3

Investigation of patch matching showed that integration of the patch extraction and patch matching procedures is necessary to ensure successful matching results. § 6.4 proposed a number of methods of increasing the number of matches found in a pair of corresponding images. These included using patches from many different feature

extraction algorithms, using patches from feature extraction algorithms with many different parameter settings, and using the features from one image to improve the extraction of features from the corresponding image. All these techniques were found to increase the number of matched patches that were detected. A further refinement to the patch extraction and patch matching procedures which was introduced but not tested was the concept of a filtering algorithm which could be used to remove patches which are likely to give false matches. As a result of the new improvements made to the combined patch extraction/patch matching procedure, the number of matches found in a pair of corresponding images was increased.

Objective 4

The success of the patch extraction and patch matching algorithms was evaluated by applying them to two pairs of small (512×512 pixels) images and one pair of large (full scene) images.

For the small images, seven matched patches were detected in one, and six matched patches were detected in the other. Thus, there were sufficient tie points to allow each pair of images to be registered, but not enough to allow a statistical analysis of the registration results. Therefore, the quality of the results had to be judged by other means. Firstly, a comparison of each pair of matched patches showed that all the patches had been correctly matched; there were no blunders in the matching. As a result, all the tie points could be considered reliable. Secondly, the spatial distribution of the tie points was considered. For one image that spatial distribution was better than the other image, meaning the accuracy of the registration would probably be better for that image. If these pair of images were being registered manually, and a similar number of tie points were used, there is no reason to suggest that the manually selected tie points would be more precise than the automatically selected tie points. If the spatial distribution of the manual tie points were similar to the spatial distribution of the automatic tie points, and the same transformation function were used in each case, then the automatic registration can be considered to be as accurate as the manual registration.

A technique based on edge matching was investigated to see if it could improve the quality of the registration results generated using patch matching. The nature of the edge matching algorithm is such that it can only be used in conjunction with the patch matching algorithm described in chapter 6. It was found that the edge matching algorithm can improve the quality of the registration results by increasing the precision

and spatial distribution of the tie points. Chapter 8 explained that where a pair of matched patches are in contact with the edge of the image it is possible that the tie points calculated from the locations of the centroids may be unreliable. However, when edges are matched instead of patches, this unreliability does not exist. In addition, since real features on the ground (the edges of features) are being used as tie points rather than arbitrary points (centroids), the edge matching tie points can be considered more precise than the patch matching tie points. For small images, edge matching was shown to give a much larger number of tie points with a much more even spatial distribution. As a result, it is reasonable to conclude that automatic registration based on patch and edge matching gives better registration results than manual registration since the automatically selected tie points are at least as precise as the manually selected tie points, but there are many more and their spatial distribution is much better. The accuracy of the automatic registration was also assessed in terms of RMS residuals. It was found that for the two tests, the RMS residuals were of the order of 1.5 pixels, depending on which edge extraction algorithms were used. It should be remembered that this value is dependent on the ability of the transformation function to represent the physical geometry of the system as well as the quality of the tie points. In order to assess the accuracy of the tie points alone, it is necessary to measure their discrepancies. However, without ground control data, this is not possible.

Chapter 8 illustrated how the patch and edge matching algorithms could be applied to full scene images. The images were split into tiles, and the algorithms were applied to selected tiles in exactly the same way as with the small images described above. Two separate tests were performed. Firstly, the full scene image were registered using just the tie points located from patch matching. A total of 39 tie points across the whole image gave a RMS residual of approximately 15 pixels. Secondly, the full scene images were registered using the tie points located from patch and edge matching. This time, 3488 tie points gave a RMS residual of just under 11 pixels. The difference between the patch matching results and the edge matching results was explained by the fact that the patch matching had used unreliable tie points, but the edge matching eliminated these. Since all the patches had been matched correctly it can be concluded that the edge matching tie points contain few blunders. Therefore, the principal reason for obtaining a RMS residual of 11 pixels must be due to the choice of transformation function, and its inability to correct for terrain induced errors. The variation of terrain height across the set of tie points was estimated to be 50m, which leads to an error of up to 7 pixels in the

across track (x) direction. Considering that the RMS residuals in x and y were 10 pixels and 4.5 pixels respectively, the results are in line with what would be expected. Although there were a large number of tie points located in the full scene images, and these tie points could be considered to be reliable, their spatial distribution was quite poor, since they were clustered into a few small regions across the image. There is no doubt that this clustering affected the accuracy of the registration, and was most likely the cause of the angular distribution of the residuals, which appeared to be systematic. The only solution to this problem is to ensure that tie points are spatially well distributed across the whole image.

Objective 5

The investigation into the automatic registration of full scene images made it clear that the affine transformation function is insufficient, and that if accurate registration is to be achieved then a more rigorous form of image registration is needed. Therefore, the idea of combining the automatically selected tie points with established photogrammetric techniques was proposed. With the availability of a DEM and a few (two or three) GCPs, it should be possible to geocode the SAR data, orthorectify the SPOT data, and register the two images to a high degree of accuracy. If ERS SAR data were used the GCPs would not be necessary. Furthermore, by implementing the proposed system, the whole process would be fully automatic.

In addition to investigating these five objectives, this thesis has also researched other aspects of the automatic image registration system, particularly preprocessing of SAR data. Chapter 4 studied speckle reduction algorithms in detail, and made some proposals for new speckle reduction techniques. It was found that these techniques were useful in this application, but further research is required to investigate how useful they will be in other circumstances.

Although the proposed method of automatic image registration has proved successful, there are a number of areas that require further research. In particular, feature extraction algorithms for both SAR and SPOT images need to be developed in order to ensure a large number of features can be matched. This is probably the most significant stumbling block in the proposed image registration system, and the one that warrants most research. Further research into the initial approximate alignment of the images would also be useful, since this is a manual processing step in the chain which could

probably be simply eliminated. The patch matching algorithm produced excellent results, but the edge matching algorithm could be improved in order to increase precision and reliability. However, the edge matching algorithm in its present form gave acceptable results, and therefore a new edge matching algorithm is not essential. Finally, further research into the use of automatically selected tie points in the rigorous geometric correction of images would be desirable. Other topics which would benefit from further research have been highlighted throughout this thesis.

This study set out to investigate automatic registration of multisensor data sets. By proposing a system of automatic image registration, researching the algorithms which make up that system, and testing the system with both small and large images, it has achieved its objectives. However, it is clear that the whole subject of automatic image registration requires more research if multisensor data sets are to be exploited to their full potential.

Chapter 10 REFERENCES

- Abbasi-Dezfouli, M. and T. G. Freeman, 1994. "Stereo-image registration based on uniform patches," *International Archives of Photogrammetry and Remote Sensing*, 30(3): 1-8.
- ASPRS, 1989. "ASPRS interim accuracy standards for large-scale maps," *Photogrammetric Engineering and Remote Sensing*, 55(7): 1038-1040.
- Barbieri, M., R. Biasutti and G. Calabresi, 1997. "Monitoring the summer 1997 floods in North-Eastern Europe through Earth Watching." *Earth Observation Quarterly*, 56-57: 6-13.
- Boardman, D., I. Dowman, A. Chamberlain, D. Fritsch and W. Newton, 1996. "An automated image registration system for SPOT data," *International Archives of Photogrammetry and Remote Sensing*, 31(4): 128-133.
- Brown, L. G., 1992. "A survey of image registration techniques," *ACM Computer Surveys*, 24(4): 325-376.
- Buongiorno, A., O. Arino, C. Zehner, P. Colagrande and P. Goryl, 1997. "ERS-2 monitors exceptional fire event in South-East Asia." *Earth Observation Quarterly*, 56-57: 1-5.
- Burrough, P., 1986. *Principles of geographical information systems for land resources assessment*. Oxford University Press, Oxford. 193 pages.
- Caesar, 1996. *Caesar SAR Image Processing Toolset User Guide*, Version 2.4a, N. A. Software Ltd. 70 pages.
- Canny, J., 1986. "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6): 679-698.

- Cook, R., I. McConnell, C. Oliver and E. Welbourne, 1994. "MUM (Merge Using Moments) segmentation for SAR images," *SAR Data Processing for Remote Sensing*, SPIE 2316: 92-103.
- Crimmins, T. R., 1986. "Geometric filter for reducing speckle," *Optical Engineering*, 25(5): 651-654.
- Curlander, J., R. Kwok and S. Pang, 1987. "A post processing system for automated rectification and registration of spaceborne SAR imagery," *International Journal of Remote Sensing*, 8(4): 621-638.
- Dai, X. and S. Khorram, 1998. "The effects of image misregistration on the accuracy of remotely sensed change detection," *IEEE Transactions on Geoscience and Remote Sensing*, 36(5): 1566-1577.
- Dare, P. M. and I. J. Dowman, 1996. "An automated procedure for registering SAR and optical imagery based on feature matching," *Microwave Sensing and Synthetic Aperture Radar*, SPIE 2958: 140-151.
- Dare, P., R. Ruskoné and I. Dowman, 1997. "Algorithm development for the automatic registration of satellite images," *Proceedings of Image Registration Workshop*, NASA Goddard Space Flight Center, Maryland. 366 pages: 83-88.
- Davis, L. S. and A. Rosenfeld, 1978. "Noise cleaning by iterated local averaging," *IEEE Transactions on Systems, Man and Cybernetics*, SMC-8: 705-710.
- Day, T. and J. P. Muller, 1989. "DEM production by stereo matching SPOT image pairs: a comparison of algorithms," *Image and Vision Computing*, 7(2): 95-101.
- Devore, R., W. Shao, J. Pierce, E. Kaymaz, B.-T. Lerner and W. Campbell, 1997. "Using non-linear wavelet compression to enhance image registration," *Wavelet Applications IV*, SPIE 3078: 539-551.
- Djamdji, J.-P., A. Bijaoui and R. Maniere, 1993. "Geometric registration of images: the multi-resolution approach," *Photogrammetric Engineering and Remote Sensing*, 59(5): 645-653.

- Dowman, I., 1998. "Automated procedures for integration of satellite images and map data for change detection: the Archangel Project." *International Archives of Photogrammetry and Remote Sensing*, 32(4): 162-169.
- Dowman, I., Z.-G. Twu, P.-H. Chen, 1997. "DEM generation from stereoscopic SAR," *In Geomatics in the Era of Radarsat*; Directorate of Geographical Operations, Canadian National Defense; CD-ROM.
- Dowman, I. J., D. Boardman and W. Newton, 1996. "A system for automatic registration of two SPOT images," *Proceedings of 22nd Annual Conference of the Remote Sensing Society*, Durham. 704 pages: 169-175.
- Einstein, A., 1905. "Zur Elektrodynamik bewegter Koerper," *Annalen der Physik*, 17: 891-921.
- ESA, 1993. *ERS User Handbook*, ESA SP-1148, European Space Agency, Noordwijk, The Netherlands.
- European Commission, 1997. *Technical Specifications for Remote Sensing Control of Surface Subsidised Arable and Forage Areas*. European Commission, Brussels.
- Flusser, J. and T. Suk, 1994. "A moment based approach to registration of images with affine geometric distortion," *IEEE Transactions on Geoscience and Remote Sensing*, 32(2): 382-387.
- Flusser, J., 1992. "An adaptive method for image registration," *Pattern Recognition*, 25(1): 45-54.
- Fonseca, L. and B. S. Manjunath, 1996. "Registration techniques for multisensor remotely sensed imagery," *Photogrammetric Engineering and Remote Sensing*, 62(9): 1049-1056.
- Fonseca, L., M. Costa, B. Manjunath and C. Kenney, 1997. "Automatic registration of satellite imagery," *Proceedings of Image Registration Workshop*, NASA Goddard Space Flight Center, Maryland. 366 pages: 13-27.
- Förstner, W. and E. Gülch, 1987. "A fast operator for detection and precise location of distinct points, corners and centres of circular features," *Proceedings of ISPRS*

Intercommission Workshop on Fast Processing of Photogrammetric Data,
Interlaken. 437 pages: 281-305.

Freeman, H., 1961. "Boundary encoding and processing," *Picture processing and psychopictorics*, Ed. B Lipkin and A. Rosenfeld, Academic Press, New York, 241-266.

Frost, V. S., J. A. Stiles, K. S. Shanmugan and J. C. Holtzman, 1982. "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4(2): 157-165.

Gonzalez, R. C. and R. E. Woods, 1992. *Digital Image Processing*. Addison-Wesley, Reading, Massachusetts. 716 pages.

Goshtasby, A., 1988. "Registration of images with geometric distortions," *IEEE Transactions on Geoscience and Remote Sensing*, 26(1): 60-64.

Goshtasby, A., G. Stockman and C. Page, 1986. "A region based approach to digital image registration with subpixel accuracy," *IEEE Transactions on Geoscience and Remote Sensing*, 24(3): 390-399.

Gredel, J. and H. Schröter, 1997. *DFD data product guide*, DLR.

Gruen, A., 1986. "High precision image matching for digital terrain model generation," *International Archives of Photogrammetry and Remote Sensing*, 26(3):284-296.

Gugan, D., 1987. "Practical aspects of topographic mapping from SPOT imagery," *Photogrammetric Record*, 12(69):349-355.

Gugan, D. and I. Dowman, 1988. "Accuracy and completeness of topographic mapping from SPOT imagery," *Photogrammetric Record*, 12 (72): 787-796.

Haan, A. de, 1991. "Contribution of the 'Politecnico di Milano' to the OEEPE test on triangulation with SPOT data," *OEEPE Official Publication* 26: 93-107.

Harris, R., 1997. "Trends in remote sensing," *Developing Space '97: Proceedings of the 1997 Remote Sensing Society Annual Student Meeting*. 129 pages: 5-6.

Hough, P. V. C., 1962. "Methods and means for recognising complex patterns, " U.S. patent 3 069 654, December 1962.

- Heipke, C., 1997. "Automation of interior, relative and absolute orientation," *ISPRS Journal of Photogrammetry and Remote Sensing*, 52: 1-19.
- Henderson, F. and A. Lewis, 1998. *Principles and applications of imaging radar*. Manual of Remote Sensing, volume 2, 3rd edition. John Wiley and Sons, New York. 866 pages.
- Kai, L. and J.-P. Muller, 1991. "Segmenting satellite imagery: a region growing scheme," *Remote sensing: global monitoring for earth management*. 11th Annual International Geoscience and Remote Sensing Symposium, Espoo, Finland. Institute of Electrical and Electronics Engineers.
- Keys, L., N. Schmidt and B. Phillips, 1990. "A prototype example of sensor fusion used for a siting analysis." *Image Processing and Remote Sensing, ACSM-ASPRS Annual Convention*, 4: 238-249.
- Kratky, V., 1989. "On-line aspects of stereophotogrammetric processing of SPOT images," *Photogrammetric Engineering and Remote Sensing*, 55(3): 311-316.
- Kuan, D. T., A. A. Sawchuk, T. C. Strand and P. Chavel, 1985. "Adaptive noise smoothing filter for images with signal dependent noise," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(2): 165-177.
- Kuan, D. T., A. A. Sawchuk, T. C. Strand and P. Chavel, 1987. "Adaptive restoration of images with speckle," *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-35(3): 373-382.
- Le Moigne, J., W. Xia, S. Chettri, T. El-Ghazawi, E. Kaymaz, B.-T. Lerner, M. Mareboyana, N. Netanyahu, J. Pierce, S. Raghavan, J. Tilton, W. Campbell and R. Crompt, 1997. "Towards an intercomparison of automated registration algorithms for multiple source remote sensing data," *Proceedings of Image Registration Workshop*, NASA Goddard Space Flight Center, Maryland. 366 pages: 307-316.
- Leberl, F., 1990. *Radargrammetric image processing*, Artech House, Norwood, USA. 595 pages.
- Lee, J. S., 1981. "Speckle analysis and smoothing of SAR images," *Computer Graphics and Image Processing* 17: 24-32.

- Lee, J. S., 1983. "A simple speckle smoothing algorithm for SAR images," *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13(1): 85-59.
- Lee, J. S., 1986. "Speckle suppression and analysis for synthetic aperture radar", *Optical Engineering*, 25(5): 636-643.
- Lewis, P., 1988. "On creating an initial segmentation of remotely sensed images as an input to an automated knowledge based segmentation system," *MSc Thesis, University of London*.
- Li, H., B. Manjunath and S. Mitra, 1995. "A contour based approach to multisensor image registration," *IEEE Transactions on Image Processing*, 4(3): 320-334.
- Light, D., 1993. "The National Aerial Photography Program as a geographic information system resource," *Photogrammetric Engineering and Remote Sensing*, 59(1): 61-65.
- Lillesand, T. and R. Kiefer, 1994. *Remote sensing and image interpretation*, 3rd edition. John Wiley and Sons, New York. 750 pages.
- Maitre, H. and Y. Wu, 1989. "A dynamic programming algorithm for elastic registration of distorted images based on autoregressive models," *IEEE Transactions on Acoustic, Speech and Signal Processing*, ASSP-37(2): 288-297.
- Marr, D. and E. Hildreth, 1980. "Theory of edge detection," *Proceedings of the Royal Society of London*, B207: 187-217.
- Massonet, D., 1997. "Satellite radar interferometry," *Scientific American*: 32-39.
- Mastin, G., 1985 "Adaptive filters for digital image noise smoothing: an evaluation," *Computer Vision, Graphics and Image Processing*, 31: 103-121.
- Mather, P., 1995. "Map-image registration accuracy using least squares polynomials," *International Journal of Geographical Information Systems*, 9(5): 543-554.
- McConnell I. and C. Oliver, 1996. "Comparison of annealing and iterated filters for speckle reduction in SAR," *Microwave Sensing and Synthetic Aperture Radar*, SPIE 2958: 74-85.

- McGlone, C., 1994. "Sensor modelling in image registration," *Proceedings of Mapping and remote sensing tools for the 21st century*, Washington DC, 110-127.
- McFarlane, N. and M. Thomas, 1984. "Speckle reduction algorithms and application to SAR images," *Proceedings of the Annual Conference of the Remote Sensing Society*, Reading. 449 pages: 391-398.
- Morgado, A. and I. Dowman, 1997. "A procedure for automatic absolute orientation using aerial photographs and a map," *ISPRS Journal of Photogrammetry and Remote Sensing*, 52: 169-182.
- Mueller, P. and R. Hoffer, 1989. "Low pass spatial filtering of Satellite radar data," *Photogrammetric Engineering and Remote Sensing*, 55(6): 887-895.
- Nagao, M. and T. Matsuyama, 1979. "Edge preserving smoothing," *Computer Graphics and Image Processing*, 9: 394-407.
- Narasimha Rao, P. V., M. S. S. R. Vidyadhar, T. Ch. Malleswara Rao and L. Venkataratnam, 1995. "An adaptive filter for speckle suppression in synthetic aperture radar images," *International Journal of Remote Sensing*, 16(5): 877-889.
- Newton, W., 1993. "An approach to the identification of forests in Thematic Mapper imagery within the context of a change detection system," *MSc Thesis, University of London*. 94 pages.
- Newton, W., C. Gurney, D. Sloggett and I. Dowman, 1994. "An approach to the automatic identification of forests and forest change in remotely sensed images," *International Archives of Photogrammetry and Remote Sensing*, 30(3): 607-614.
- Oliver, C., D. Blacknell, I. McConnell and R. White, 1995. "Optimum edge detection in SAR," *Synthetic aperture radar and passive microwave sensing*, SPIE 2584: 152-163.
- O'Neill, M., 1991. "A dynamic numerical camera model for the SPOT-1 sensor," *Ph.D. Thesis, University of London*. 475 pages.

- Orti, F., 1981. "Optimal distribution of tie points to minimize Landsat image registration errors," *Photogrammetric Engineering and Remote Sensing*, 47: 101-110.
- Otto, G. and T. Chau, 1989. "A region growing algorithm for matching terrain images," *Image and Vision Computing*, 7(2): 83-94.
- Peet, F. G., 1975. "The use of invariants in the transformation and registration of images," *The Canadian Surveyor*, 29(5): 501-506.
- Pohl, C., 1996. "Geometric aspects of multisensor image fusion for topographic map updating in the humid tropics." *Ph.D. Thesis, ITC publication number 39*. 159 pages.
- Pohl, C. and J. L. van Genderen, 1998. "Multisensor image fusion in remote sensing: concepts, methods and applications." *International Journal of Remote Sensing*, 19(5): 823-854.
- Pratt, W. K., 1991, *Digital Image Processing*, 2nd edition. John Wiley, New York. 698 pages.
- Prechtel, N and O. Bringmann, 1998. "Near real time road extraction from satellite images using vector reference data." *International Archives of Photogrammetry and Remote Sensing*, 32(2): 229-234.
- Renouard, L. and F. Perlant, 1993. "Geocoding SPOT products with ERS-1 geometry," *Proceedings of the Second ERS-1, Space at the Service of the Environment*, Hamburg: 653-658.
- Rignot, E., R. Kwok, J. Curlander and S. Pang., 1991. "Automated multisensor registration: requirements and techniques," *Photogrammetric Engineering and Remote Sensing*, 57(8): 1029-1038.
- Rogers, R. and L. Wood, 1990. "The history and status of merging multiple sensor data: an overview." *Image Processing and Remote Sensing, ACSM-ASPRS Annual Convention*, 4: 352-360.

- Ruskoné, R. and I. Dowman, 1997. "Segmentation design for an automatic multisource registration," *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III*, SPIE 3072: 307-317.
- Satter, M. and A. Goshtasby, 1997. "Registration of deformed images," *Proceedings of Image Registration Workshop*, NASA Goddard Space Flight Center, Maryland. 366 pages: 221-231.
- Schistad-Solberg, A., A. Jain and T. Taxt, 1994. "Multisource classification of remotely sensed data: fusion of Landsat TM and SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, 32(4): 768-778.
- Schrier, G., 1993. *SAR Geocoding: Data and Systems*. Wichman Verlag, Karlsruhe, Germany. 435 pages.
- Sowter, A., 1998. "The automatic extraction of DEM data from stereo Radarsat pairs over the tropics," *International Archives of Photogrammetry and Remote Sensing*, 32(2): 291-298.
- SPOT, 1998. SPOT Image Co. internet document: www.spotimage.fr.
- Stockman, G., S. Kopstein and S. Benett, 1982. "Matching images to models for registration and object detection via clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(3): 229-241.
- Toll, D., 1985 "Analysis of digital LANDSAT MSS and SEASAT SAR data for use in discriminating landcover at the urban fringe of Denver, Colorado," *International Journal of Remote Sensing* 6(7): 1209-1229.
- Toutin, T., 1995. "Intégration des données multisources: comparaison de methodes géométriques et radiométriques," *International Journal of Remote Sensing*, 16(15): 2795-2811.
- Tseng, Y., J. Tzen, K. Tang and S. Lin, 1997. "Image to image registration by matching area features using Fourier descriptors and neural networks," *Photogrammetric Engineering and Remote Sensing*, 63(8): 975-983.
- Twu, Z.-G., 1996. "Automatic height extraction from stereoscopic SAR imagery" *Ph.D. thesis, University of London*. 222 pages.

- Ventura, A., A. Rampini and R. Schettini., 1990. "Image registration by recognition of corresponding structures," *IEEE Transactions on Geoscience and Remote Sensing*, 28(3): 305-314.
- Vohra, V. and I. Dowman, 1997. "Automatic extraction of significant features from high resolution satellite images for registration with a map," *Developing Space '97: Proceedings of the 1997 Remote Sensing Society Annual Student Meeting*, London. 126 pages: 23-29.
- Vohra, V. and I. Dowman, 1996. "Automated extraction of large buildings from high resolution satellite images for registration with a map," *International Archives of Photogrammetry and Remote Sensing*, 31(3): 903-908.
- Westin, T., 1990. "Precision rectification of SPOT imagery," *Photogrammetric Engineering and Remote Sensing*, 55(3): 311-316.
- Weszka, J., 1978. "A survey of threshold selection techniques," *Computer Graphics and Image Processing*, 7: 259-265.
- White, R., 1991. "Change detection in SAR imagery," *International Journal of Remote Sensing*, 12(2): 339-360.
- Wolf, P., 1983. *Elements of Photogrammetry* 2nd edition. McGraw-Hill, New York. 628 pages.