

DSP

DIGITAL SIGNAL AND IMAGE PROCESSING SERIES



Processing of Synthetic Aperture Radar Images

Edited by Henri Maître

ISTE

WILEY

Processing of Synthetic Aperture Radar Images

Edited by
Henri Maître



This page intentionally left blank

Processing of Synthetic Aperture Radar Images

This page intentionally left blank

Processing of Synthetic Aperture Radar Images

Edited by
Henri Maître



First published in France in 2001 by Hermes Science Publications entitled "Traitement des images de RSO"

First published in Great Britain and the United States in 2008 by ISTE Ltd and John Wiley & Sons, Inc.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced, stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers, or in the case of reprographic reproduction in accordance with the terms and licenses issued by the CLA. Enquiries concerning reproduction outside these terms should be sent to the publishers at the undermentioned address:

ISTE Ltd
6 Fitzroy Square
London W1T 5DX
UK
www.iste.co.uk

John Wiley & Sons, Inc.
111 River Street
Hoboken, NJ 07030
USA
www.wiley.com

© ISTE Ltd, 2008
© Hermes Science, 2001

The rights of Henri Maître to be identified as the author of this work have been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Library of Congress Cataloging-in-Publication Data

[Traitement des images de RSO. English] Processing of synthetic aperture radar images /
edited by Henri Maître.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-84821-024-0

1. Synthetic aperture radar. I. Maître, Henri.

TK6592.S95T73 2008

621.3848--dc22

2007022559

British Library Cataloguing-in-Publication Data

A CIP record for this book is available from the British Library

ISBN: 978-1-84821-024-0

Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire.



Cert no. SGS-COC-2933

www.fsc.org

© 1996 Forest Stewardship Council

Table of Contents

Introduction	xv
Chapter 1. The Physical Basis of Synthetic Aperture Radar Imagery	1
Jean-Marie NICOLAS and Sylvie LE HÉGARAT-MASCLE	
1.1. Electromagnetic propagation	1
1.1.1. The laws of propagation in homogenous media	1
1.1.1.1. Basic equations	1
1.1.1.2. Propagation equation solution	3
1.1.2. Propagation equation in heterogenous media	5
1.1.2.1. The permittivity variation case	6
1.1.2.2. The propagation equation in an absorbing medium	8
1.1.3. Application to satellite radars	8
1.2. Matter-radiation interaction	10
1.2.1. Theoretical backscattering models	10
1.2.1.1. Unbounded flat interface	10
1.2.1.2. Rayleigh scatterer	12
1.2.2. Phenomenological backscattering models	12
1.2.2.1. Rough interface	12
1.2.2.2. Scattering by a generic target	13
1.2.2.3. Scattering by a set of targets	13
1.3. Polarization	14
1.3.1. Definitions	14
1.3.2. Wave polarization	16
1.3.3. The BSA convention	18
1.3.4. Complex backscattering matrix S , Mueller matrix	19
1.3.4.1. Properties of M	22
1.3.4.2. Other definitions of M	23

Chapter 2. The Principles of Synthetic Aperture Radar	25
Jean-Marie NICOLAS and FrédéricADRAGNA	
2.1. The principles of radar	26
2.1.1. Description of surveillance radars	26
2.1.2. Notion of resolution	29
2.1.3. Imaging radars	30
2.1.4. Synthetic Aperture Radar (SAR)	32
2.1.5. The radar equation	33
2.2. The SAR equations	35
2.2.1. The principle of pulse compression	35
2.2.2. Range resolution	38
2.2.3. Azimuth resolution	41
2.2.3.1. Principles	41
2.2.3.2. Application to the ERS-1 satellite	43
2.2.3.3. Taking platform motions into account	44
2.2.4. Transfer function of the SAR antenna	45
2.2.5. Range and azimuth samplings	46
2.2.5.1. Range sampling	46
2.2.5.2. Azimuth sampling	46
2.2.5.3. Choice of a PRF: theoretical constraints	47
2.2.6. Operation of a satellite SAR	47
2.2.6.1. Effects on radial quality	48
2.2.6.2. Effects on azimuth quality	49
2.2.6.3. The squint case	49
2.3. Acquisition geometry of SAR images	50
2.3.1. Sampling and resolution	51
2.3.1.1. Range	51
2.3.1.2. Azimuth	52
2.3.2. Radar geometry and ground geometry	53
2.3.2.1. Radar geometry	53
2.3.2.2. Ground geometry	53
2.3.3. Overlays, foreshortening and shadow	54
Chapter 3. Existing Satellite SAR Systems	57
Jean-Marie NICOLAS and FrédéricADRAGNA	
3.1. Elements of orbitography	57
3.1.1. Remote sensing satellite orbits	58
3.1.1.1. Inclination	58
3.1.1.2. Period	58
3.1.1.3. Eccentricity	60
3.1.1.4. Sun-synchronism	60
3.1.1.5. Cycle	61
3.1.1.6. Phasing	61
3.1.1.7. Orbital drift and correction	61

3.2. Polar orbiting SAR satellites	62
3.2.1. SEASAT	62
3.2.2. ERS	62
3.2.3. JERS	64
3.2.4. RADARSAT	64
3.3. Satellites in non-polar orbit	65
3.3.1. ALMAZ	65
3.3.2. Space shuttle use	66
3.3.2.1. Columbia: the SIR-A program (1982)	67
3.3.2.2. Challenger: the SIR-B program (1984)	67
3.3.2.3. Endeavour: the SIR-C program (1994)	68
3.3.2.4. Endeavour: the SRTM program (2000)	69
3.4. Other systems	69
3.5. Airborne SARs	70
3.5.1. Sandia	70
3.5.2. CCRS	70
3.5.3. CRL/NASDA	71
3.5.4. JPL's AIRSAR	71
3.5.5. PHARUS	72
3.5.6. DLR's E-SAR	72
3.5.7. Some remarkable military systems	72
Chapter 4. Synthetic Aperture Radar Images	75
Frédéric ADRAGNA, Sylvie LE HÉGARAT-MASCLE and Jean-Marie NICOLAS	
4.1. Image data	76
4.1.1. Raw data	76
4.1.2. Complex single-look data	76
4.1.3. Multi-look data	78
4.1.4. Derivative products	79
4.1.5. Polarimetric data	80
4.1.5.1. SIR-C/X-SAR data	80
4.1.5.2. AirSAR data	81
4.1.5.3. Polarimetric synthesis and examples	82
4.2. Radiometric calibration	83
4.2.1. Scalar SAR data calibration	83
4.2.2. Calibration of SAR polarimetric data	84
4.3. Localization precision	85
Chapter 5. Speckle Models	87
Armand LOPÈS, René GARELLO and Sylvie LE HÉGARAT-MASCLE	
5.1. Introduction to speckle	87
5.2. Complex circular Gaussian model of single-look and scalar speckle	89
5.2.1. Random walk and Gaussian speckle conditions	90

5.2.1.1. Random walk in a complex plane	90
5.2.1.2. Fully developed speckle hypotheses	91
5.2.2. First order distributions: R radar reflectivity	92
5.2.2.1. Fully developed complex speckle distribution: radar reflectivity	92
5.2.2.2. Distributions of amplitude A, phase ϕ and intensity I	94
5.2.2.3. Multiplicative and additive speckle models	96
5.3. Complex circular multi-variate Gaussian model of vectorial or multi-look speckle	98
5.3.1. Joint distribution of single-look complex data	98
5.3.2. Spatial correlation and spectral properties of the single-look Gaussian speckle	99
5.3.3. Joint distribution of single-look detected data	104
5.3.4. Distributions of scalar multi-look data	105
5.3.4.1. Gamma and generalized Gamma distributions	107
5.3.4.2. Goodman distribution	110
5.3.4.3. Multi-look logarithmic data	112
5.3.4.4. Spectral and spatial properties of multi-look speckle	115
5.3.5. Multi-look vectorial distributions	115
5.3.5.1. Distribution of the variables derived from Σ_z and the Wishart distribution	116
5.4. Non-Gaussian speckle models	125
5.4.1. Scalar product model and normalized texture	126
5.4.1.1. Normalized texture variable	126
5.4.1.2. Statistical properties of heterogenous area data	127
5.4.1.3. The Pearson system and K, U, B and W distributions	129
5.4.2. Non-Gaussian clutter	133
5.4.2.1. The lognormal distribution	134
5.4.2.2. Weibull distribution	134
5.4.2.3. Nagakami-Rice distribution	135
5.4.2.4. The partially developed speckle	135
5.5. Polarimetric radar speckle	136
5.5.1. The Gaussian model	136
5.5.2. The product model	137
5.5.3. Distributions of the phase difference and 2-amplitude ratio	138
5.5.4. Correlation of the co- and cross-polar channels	140
Chapter 6. Reflectivity Estimation and SAR Image Filtering	143
Armand LOPÈS, Florence TUPIN and Sylvie LE HÉGARAT-MASCLE	
6.1. Introduction	143
6.2. Estimations of reflectivity R	145
6.2.1. Bayesian estimation recap	145
6.2.1.1. The Kronecker function: $L(x, x') = 1 - \delta_{x, x'}$	146
6.2.1.2. The quadratic cost function: $L(x, x') = (x - x')^2$	146

6.2.1.3. Notes on estimator behavior	147
6.2.2. Estimation of constant reflectivity R	148
6.2.2.1. Multi-look filters	148
6.2.2.2. Spatial filters	150
6.3. Single-channel filters with <i>a priori</i> knowledge of the scene	151
6.3.1. Linear MMSE filters	151
6.3.1.1. Lee and Kuan filters	151
6.3.1.2. Frost filter	152
6.3.2. Non-linear MMSE filters	154
6.3.2.1. T-linear filters	154
6.3.2.2. APE filters	155
6.3.3. Non-linear MAP filters	156
6.4. Multi-channel filters	157
6.4.1. Intensity data	157
6.4.1.1. Homogenous channels	157
6.4.1.2. Textured channels	158
6.4.2. A general case: the linear vector MMSE	159
6.5. Polarimetric data filtering	161
6.5.1. Speckle reduction filters based on a combination of polarimetric channels	161
6.5.2. Polarimetric data restoration filters	165
6.6. Estimation of filter parameters	167
6.6.1. Formulae for statistical estimation	167
6.6.2. Choice of a calculation window	168
6.7. Filter specificities	169
6.7.1. Statistical evaluation of the various estimators	169
6.7.2. Results on an ERS-1 image	171
6.8. Conclusion	172
Chapter 7. Classification of SAR Images	175
Danielle DUCROT, Florence TUPIN and Sylvie LE HÉGARAT-MASCLE	
7.1. Notations	176
7.2. Bayesian methods applied to scalar images	177
7.2.1. Pixel-based Bayesian methods	179
7.2.1.1. Supervised methods	179
7.2.1.2. Unsupervised methods	180
7.2.2. Contextual Bayesian methods	181
7.2.2.1. Supervised methods	182
7.2.2.2. Unsupervised methods	184
7.2.3. Global Bayesian methods	185
7.2.3.1. Supervised methods	185
7.2.3.2. Unsupervised methods	186
7.3. Application of the Bayesian methods to ERS-1 time series	186
7.3.1. Pixel classification	188

7.3.1.1. The single-channel case	188
7.3.1.2. Multi-temporal contribution	188
7.3.1.3. Introduction of derived channels	189
7.3.1.4. Contribution of contextual classification methods	189
7.3.2. Radar and SPOT fusion, comparison with SPOT	191
7.3.3. Conclusion	192
7.4. Classification of polarimetric images	192
7.4.1. Classification in three major backscattering mechanisms	193
7.4.1.1. The principle	193
7.4.1.2. Implementation and illustration of results	193
7.4.1.3. Conclusion	195
7.4.2. Entropy-based classification	195
7.4.2.1. The principle	195
7.4.2.2. Implementation and illustration of results	196
7.4.3. Polarimetric Bayesian classifications	198
7.4.3.1. The single-look Gaussian case	198
7.4.3.2. The multi-look Gaussian case	200
7.4.3.3. The product model case	201
7.4.4. Polarimetric non-Bayesian classifications	201
Chapter 8. Detection of Points, Contours and Lines	207
Florence TUPIN and Armand LOPÈS	
8.1. Target detectors	208
8.1.1. Intensity threshold	208
8.1.2. Ratio use	210
8.1.3. Other approaches	213
8.1.4. Conclusion	213
8.2. Contour detectors	214
8.2.1. Average ratio	215
8.2.2. Likelihood ratio	218
8.2.2.1. Use of detected data	218
8.2.2.2. Complex data use	219
8.2.3. Exponentially-balanced average ratio	221
8.2.4. Multi-directional case	225
8.3. Line detectors	225
8.3.1. Average ratio	226
8.3.2. Correlation with an ideal structure	227
8.4. Line and contour connection	228
8.5. Conclusion	231
Chapter 9. Geometry and Relief	233
Frédéric ADRAGNA	
9.1. Radar image localization	233
9.1.1. Precision of orbital data	234

9.1.2. Distance localization	234
9.1.3. Azimuth localization	235
9.1.4. Three-dimensional point localization from radar imaging	237
9.1.5. Moving targets	238
9.2. Geometric corrections	239
9.2.1. Geometric corrections with DTM	239
9.2.2. Image registration	240
9.2.2.1. Case of images obtained from identical orbits	240
9.2.2.2. Images acquired from different incidences	241
9.2.2.3. Images obtained in opposite direction orbits (ascending/descending)	241
Chapter 10. Radargrammetry	243
Henri MAÎTRE	
10.1. Stereovision principles: photogrammetry	243
10.1.1. Principles	243
10.1.2. Stereovision methods	244
10.2. Principles of radargrammetry	245
10.2.1. Lateral vision effects	246
10.2.2. Geometric reconstruction	249
10.2.2.1. Disparity/altitude relation in the general case	249
10.2.2.2. Simplified disparity/altitude relations	250
10.2.3. Matching	251
10.2.3.1. Correlation	251
10.2.3.2. Pyramidal approach	254
10.2.4. Geometry effects connected with radar viewing	255
10.2.5. Filling of hidden parts and non-matched areas	257
10.3. Results of radargrammetry	257
10.4. Conclusion	259
Chapter 11. Radarclinometry	261
Henri MAÎTRE	
11.1. Radarclinometry equation	262
11.1.1. Shape from shading	262
11.1.2. Shape from shading and radar	263
11.1.3. Radarclinometry equation	263
11.2. Resolution of the radarclinometry equation	265
11.2.1. Wildey approach	265
11.2.2. Frankot and Chellappa approach	266
11.2.3. Resolution of radargrammetry equation without <i>a priori</i> knowledge	267
11.2.3.1. Inversion of equation [11.6]	268
11.2.3.2. Regularization	268

11.3. Determination of unknown parameters	271
11.3.1. Determination of σ^0	271
11.3.2. Determination of backscattering distribution	272
11.4. Results	273
11.5. Radarpolarimetry	276
11.6. Conclusion	278
Chapter 12. Interferometry	279
Frédéric ADRAGNA and Jean-Marie NICOLAS	
12.1. Interferometry principle	280
12.1.1. Phase and interference	280
12.1.2. Construction of an interferometric product	281
12.1.3. What type of information does the interferogram carry?	282
12.2. Interferogram modeling	283
12.2.1. Statistical laws of interferograms	283
12.2.2. Estimation of empirical coherence	284
12.3. Geometric analysis of data	285
12.4. Applications of interferometry	288
12.4.1. Elevation measurement	288
12.4.1.1. Limitations	290
12.4.1.2. Precisions	291
12.4.2. Land movement measurement: differential interferometry	291
12.4.3. Use of coherence	292
12.4.3.1. Implementation	293
12.5. Limitations of interferometry imaging	294
12.5.1. Temporal change limitations	294
12.5.2. Limitations due to geometry	294
12.5.3. Effects of atmospheric propagation	297
12.5.3.1. Troposphere case	298
12.5.3.2. Ionosphere case	299
12.5.4. Role of signal-to-noise ratio	299
Chapter 13. Phase Unwrapping	301
Emmanuel TROUVÉ and Henri MAÎTRE	
13.1. Introduction	301
13.1.1. Corrected phase shift	302
13.1.2. Residue detection	304
13.2. Preprocessing of InSAR data	305
13.2.1. Analysis of local frequencies	306
13.2.1.1. Modeling	306
13.2.1.2. Direct estimation method	307
13.2.1.3. Confidence measure	309
13.2.1.4. Multi-scale implementation	310
13.2.2. Interferogram filtering	312

13.2.2.1. Two different approaches	312
13.2.2.2. Fringe pattern compensation	313
13.2.3. Detection of non-unwrappable areas	317
13.3. Phase unwrapping methods.	319
13.3.1. Local methods	320
13.3.1.1. Placing the cuts	321
13.3.1.2. Finding a mask	321
13.3.2. Least squares.	322
13.3.2.1. Considering disturbances	323
13.3.2.2. Minimization algorithms.	325
13.3.3. Green's functions	326
13.3.4. Conclusion	329
Chapter 14. Radar Oceanography	331
Jean-Marc LE CAILLEC and René GARELLO	
14.1. Introduction to radar oceanography	331
14.2. Sea surface description	334
14.2.1. Division of the sea spectrum into different scales	334
14.2.2. Reflection of electromagnetic waves from the sea surface	336
14.3. Image of a sea surface by a real aperture radar	337
14.3.1. Tilt modulation	337
14.3.2. Hydrodynamic modulation	341
14.4. Sea surface motions	344
14.5. SAR image of the sea surface	345
14.5.1. Quasi-linear approximation of the SAR spectrum image	345
14.5.2. Modeling of non-linearities	348
14.6. Inversion of the SAR imaging mechanism.	352
14.6.1. SAR spectrum inversion	352
14.6.2. Underwater topography inversion	353
Bibliography	357
List of Authors	377
Index	379

This page intentionally left blank

Introduction

Synthetic aperture radar imagery was born as a result of an exciting process extending over more than half a century and involving parallel advances in physics, electronics, signal processing and finally image processing. As radars first appeared at the eve of World War II, their prime task was surveillance, i.e. detection. They gradually acquired reconnaissance capabilities: very low resolution images were produced by space scans, while a persistent display made it possible to distinguish echoes from different reflectors. To move on to an actual image, all that was needed was to accelerate the scan and organize systematically collected echoes along two directions. But above all, major improvements had to be achieved on two key parameters: resolution, which, due to the particular wavelengths that were used, was rather poor at a useful monitoring range, and on the other hand, discriminating power, i.e., receiver sensitivity to major relevant dynamics. Both parameters were improved in the wake of manifold technical progress, but also thanks to some decisive choices, including side vision that helped remove the dominant echo of orthogonal reflection and synthetic aperture that paved the way to virtually unlimited resolution capabilities. As uncomplicated as these ideas may appear, they could not have materialized without proper technological backing. It thus took a continuous movement back and forth between methodological, conceptual strides and progress in areas such as sensors and emitters, electronic components and processing algorithms for radar imaging to eventually emerge on a par with optical imaging as a basic remote sensing tool.

By the 1960s, the essentials of radar imaging that make it so attractive nowadays had been investigated and recorded. Its foundations ranging from the capacity of discriminating among different materials to that of penetrating through various covers and vegetation layers, from geometrical effects to depolarization properties, from stereoscopic, interferometric and clinometric capacities to differential wavelength properties, had all been laid down. This progress, however, was not widely publicized. Born on the spur of anti-aircraft defense needs, radar imaging

was still closely connected with military applications. As a result, even its most outstanding advances, which were often regarded as strategic, were very slow to seep into other areas of industry or research. By its very complexity, particularly its hi-tech requirements, radar imaging was out of bounds for many industrial applications, and academics would not get into it without solid support from some powerful constructors. Even having a look at images from synthetic aperture radars was a lot of trouble. This was not only due to obvious property restrictions, but also to the complex way in which they were obtained. These images, which often were the product of experimental sensors, were very hard to use. The basic acquisition parameters that will be detailed further in this work were subject to endless adjustments. Intermediate processing was also constantly improving, involving transient changes that were not always fully documented. To users, a significant leap forward was made with the advent of civilian satellite sensors such as SEASAT, SIR-A and -B, and especially the ERS family. These systems made it possible to establish a number of reference products that became accessible to all laboratories and helped expand the application range to a considerable extent. Whole areas, from natural disaster prevention to geological and mining surveys, from cartography to polar route monitoring and from forestry management to sea surveys, thus opened up to the use of radar imaging.

Radar imaging has numerous advantages over optical imaging. From among them, we have to underline its capacity of working in any weather, which is particularly useful in frequently overcast countries such as those located in the equatorial belt. In addition, its coherent imaging properties (i.e., its capacity of collecting amplitude and phase signals) are used to attain remarkable resolutions in the synthetic aperture version, while interferometry uses them to measure extremely fine altitudes and control some even finer displacements (accounting for bare fractions of the operating wavelength). The penetration capacity of radar waves is also linked to microwave frequency. It helps them get across light foliage and detect underground structures provided they are shallowly buried in very dry environments. Finally, radar waves are for the most part polarized, and the extent to which they are depolarized by different media that backscatter them is a great source of information for agriculture, geology and land management.

Nevertheless, radar imaging is less attractive for its edge over optical imaging than for the way it complements the latter. The formation of radar images, for instance, is governed by *time-of-flight* laws rather than the *projection* imaging we are familiar with. Moreover, radar imaging is especially sensitive to the geometric properties of targets, whether microscopic (e.g., roughness, surface effects) or macroscopic (e.g., orientation, multiple reflections). On the other hand, optical imaging is more sensitive to the physicochemical properties (e.g., emissivity, albedo, color) of targets. Radars are sensitive to properties such as the *nature* of materials (metallic targets, for example) and their condition (such as soil humidity or

vegetation dryness) that optics is frequently unable to perceive. Finally, optical imaging depends on a source of light, which is usually the Sun, while radar imaging has nothing to do with this. As a result, radar images, as compared to optical images, have higher daytime and seasonal stability, but depend to a greater extent on the position of the sensor when it takes a shot.

For all these reasons, many satellites have been equipped with imaging radars. While some of them were merely experimental, others lived on through their descendants, such as the Lacrosse family, which are US military satellites, and the ERS, which are European civilian satellites. These satellites are permanent sources of information on our planet. Such information is mostly processed by photograph interpreters, but automatic techniques are gaining ground, driven by an increased amount of images that need to be processed and the growing demand for reliable and quantitative measurements. This work is designed to contribute to the development of such automatic methods.

It therefore covers the three basic types of tools that are required to digitally process images supplied by synthetic aperture radars, namely:

- physical concepts that help account for the main propagation phenomena and substance-radiation interactions and provide notions on how radars and their supporting platforms operate;
- mathematical models that statistically describe the very peculiar characteristics of radar issued signals and the properties we may expect of them; and
- image processing methods to suit specific applications: detection, reconnaissance, classification or interpretation.

The careful, simultaneous consideration of these three types of properties has helped devise effective automatic methods for extracting information from radar images. For many years, users only adapted to radar imaging a number of algorithms that had worked successfully in optical remote sensing. These commercially available programs were well known to photo interpreters and had been amply tested on images from Landsat, SPOT and Meteosat. Applied to radar images, they yielded very poor results, strengthening the belief that these were definitely unmanageable by automatic tools and open to nothing but qualitative interpretation. It is the goal of this work to disprove this belief and provide the necessary elements for turning the asset of radar imagery to good account.

This page intentionally left blank

Chapter 1

The Physical Basis of Synthetic Aperture Radar Imagery

1.1. Electromagnetic propagation

The physics behind radar image formation is complex and involves several different topics. Some deal with electronic components devoted to transmission and reception of the wave, but they will not be discussed here. Other aspects, namely wave propagation and the interaction between microwave frequency waves and materials, are more important for our purposes. These two topics are the subject of this chapter. Electromagnetism obviously underlies both these phenomena and we begin with a review of useful results in this area.

1.1.1. *The laws of propagation in homogenous media*

1.1.1.1. *Basic equations*

An electromagnetic wave such as that emitted by radars is characterized at any point in space and at every moment by four vector values: \vec{E} (electric field), \vec{D} (electric displacement), \vec{B} (magnetic induction) and \vec{H} (magnetic field).

2 Processing of Synthetic Aperture Radar Images

These quantities verify Maxwell's equations, which in the absence of free charges and current densities are written as [JAC 75]:

$$\begin{aligned}\operatorname{div} \vec{\mathbf{D}} &= 0, & \vec{\mathbf{rot}} \vec{\mathbf{H}} &= \frac{\partial \vec{\mathbf{D}}}{\partial t}, \\ \operatorname{div} \vec{\mathbf{B}} &= 0, & \vec{\mathbf{rot}} \vec{\mathbf{E}} &= -\frac{\partial \vec{\mathbf{B}}}{\partial t}.\end{aligned}$$

In the linear stationary case, the fields, the electric displacement and the magnetic induction are ruled by the following relations:

$$\vec{\mathbf{D}} = \epsilon \vec{\mathbf{E}}, \quad \text{and} \quad \vec{\mathbf{B}} = \mu \vec{\mathbf{H}}.$$

where ϵ is the permittivity and μ is the permeability. We will consider them as scalar values in this book (they are tensors in the general case of anisotropic dielectrics).

The electric field $\vec{\mathbf{E}}$ and magnetic field $\vec{\mathbf{H}}$ vectors are sufficient to characterize this electromagnetic wave for an unbounded, homogenous, isotropic medium which is free of charges and currents. We will use Maxwell's equations to show that every component of these fields verifies the wave equation:

$$\nabla^2 u - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad \text{where} \quad v = \frac{1}{\sqrt{\epsilon \mu}} \quad [1.1]$$

We thus observe the electromagnetic energy transmission; v is the propagation velocity of the electromagnetic wave.

By denoting ϵ_0 the vacuum permittivity and μ_0 the vacuum permeability, we deduce c , i.e. the speed of light, as being:

$$c = \frac{1}{\sqrt{\epsilon_0 \mu_0}}$$

In the general case and in the absence of any charge or current, the relative permittivity $\epsilon_r = \frac{\epsilon}{\epsilon_0}$ and relative permeability $\mu_r = \frac{\mu}{\mu_0}$ of the propagation medium are normally used, which makes it possible to express the propagation velocity according to c :

$$v = \frac{c}{\sqrt{\epsilon_r \mu_r}}$$

The refractive index n for a propagation medium is defined as:

$$n = \frac{c}{v} = \sqrt{\epsilon_r \mu_r}.$$

Note that in non-magnetic media we will deal with $\mu_r = 1$, which leads to $n = \sqrt{\epsilon_r}$.

Since the medium is unbounded, $\vec{E}(\vec{r})$ and $\vec{H}(\vec{r})$ are perpendicular to each other at any \vec{r} , and both are perpendicular to the propagation direction $\vec{s}(\vec{r})$ that represents the energy path, which is also called a ray.

If a preferred direction can be specified by convention in the plan (\vec{E}, \vec{H}) , we will then be able to characterize \vec{E} (and therefore \vec{H}) in terms of its polarization, i.e., its orientation with respect to the defined direction.

1.1.1.2. Propagation equation solution

In the presence of an isotropic radiation source $g(\vec{r}_0, t)$ located at \vec{r}_0 , the solution of propagation equation [1.1] at any \vec{r} point in space is written:

$$u(\vec{r}, t) = \frac{1}{4\pi |\vec{r} - \vec{r}_0|} g\left(\vec{r}_0, t - \frac{|\vec{r} - \vec{r}_0|}{v}\right) \quad [1.2]$$

The wave then propagates from the source (homogenous medium) in such a way that the wavefront, i.e., the normal ray surface everywhere in space, is a sphere

4 Processing of Synthetic Aperture Radar Images

centered on the source: the propagation between the source and any observer is carried out in a straight line.

In the specific case of satellite systems, the objects impinged by the electromagnetic wave are far enough from the antenna so that the wave can be seen as locally plane around the study zone (Fraunhofer zone). Moreover, the only waves that are generally taken into account are quasi-monochromatic waves that have a frequency f_c (harmonic case) and are defined by their wavelength $\lambda = \frac{v}{f_c}$ and wave vector $\vec{k} = \frac{2\pi}{\lambda} \vec{s}$.

Given these hypotheses, we show that in the presence of a source in \vec{r}_0 , u , which is the propagation equation solution in \vec{r} , is written as:

$$u(\vec{r}, t) = \frac{e^{i\vec{k}(\vec{r}-\vec{r}_0)}}{4\pi|\vec{r}-\vec{r}_0|} g(\vec{r}_0, t) \quad [1.3]$$

\vec{r} and \vec{r}_0 fields differ from one another by a phase term and an attenuation term (a term in $\frac{1}{|\vec{r}-\vec{r}_0|}$). A surface defined by a set of points sharing the same phase is called a wave surface: \vec{k} is normal to the wave surface at every point, and the electric and magnetic fields are situated in the plan tangent to the wave surface.

In the general case (equation [1.2]) as well as in the quasi-monochromatic case (equation [1.3]), the term $\frac{1}{|\vec{r}-\vec{r}_0|}$ appears describing an attenuation phenomenon arising from energy conservation. By integrating energy over a wave surface or the wave front, energy transmitted by the source should be obtained. This attenuation effect, which is quite strong in airborne radars, may also be significant in satellite imaging radars. With the transmitter located hundreds of kilometers away in orbit and imaged areas extending over dozens of kilometers, the attenuation term may indeed vary by several percentage points and create noticeable effects on the images.

1.1.2. Propagation equation in heterogenous media

As the wave no longer propagates through a homogenous medium, electric and magnetic fields no longer obey propagation equation [1.1]. In this case, several major phenomena have to be taken into account, i.e.:

- a change in the propagation, which is no longer a straight line due to a wavefront bend;
- a scattering phenomenon (e.g., backscattering, multiple scattering) that alters the energy transmitted along a ray;
- a potential transfer of energy into heat leading to wave absorption.

As a general rule, a simple expression of the propagation equation will no longer be available. Nevertheless, if the perturbation caused by the heterogeneities of the propagation medium is weak enough, we can resort to a traditional method (also in the linear framework) which consists of adding a complementary term to equation [1.1]:

$$\nabla^2 u - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2} = \Lambda(u, \vec{r}) \quad [1.4]$$

where

$$\Lambda(u, \vec{r}) = \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} - \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}$$

By decomposing the field u into two terms: u_0 which is the incident field and u_p which is the field created by the effects of the perturbation:

$$u = u_0 + u_p \quad u_0 \text{ such that } \nabla^2 u_0 - \frac{1}{c^2} \frac{\partial^2 u_0}{\partial t^2} = 0,$$

we note that, in the end, the problem comes down to solve the following equation:

$$\nabla^2 u_p - \frac{1}{c^2} \frac{\partial^2 u_p}{\partial t^2} = \Lambda(u_0 + u_p, \vec{r})$$

6 Processing of Synthetic Aperture Radar Images

If there is no absorption, we can use the Born first order approximation to solve it by only taking into consideration the incident field within the perturbation term:

$$\nabla^2 u_p - \frac{1}{c^2} \frac{\partial^2 u_p}{\partial t^2} = \Lambda(u_0, \vec{r}),$$

This makes it possible to interpret the term $\Lambda(u_0, \vec{r})$ as a source term in the field u_p , thus explaining the wave scattering process.

1.1.2.1. The permittivity variation case

In the case of permittivity variations, we can show that the propagation equation verified by field $\vec{E}(\vec{r})$ is in general written as [LAV 97]:

$$\nabla^2 \vec{E} - \frac{1}{v^2(\vec{r})} \frac{\partial^2 \vec{E}}{\partial t^2} = -\nabla(\vec{E} \cdot \nabla(\text{Log}\epsilon_r)). \quad [1.5]$$

In the harmonic case, the second term of this equation can be ignored if the permittivity variations verify the relation:

$$|\nabla(\text{Log}\epsilon_r)| \ll \frac{1}{\lambda}, \quad [1.6]$$

Taking into account a variation $\Delta\epsilon$ over a displacement Δr , the above relation can also be written as:

$$\frac{\Delta\epsilon}{\epsilon} \ll \frac{\Delta r}{\lambda}.$$

Under these assumptions, the propagation equation is:

$$\nabla^2 \vec{E} - \frac{1}{v^2(\vec{r})} \frac{\partial^2 \vec{E}}{\partial t^2} = 0. \quad [1.7]$$

This occurs as if we had replaced v by $v(\vec{r})$.

Two cases have to be considered:

– when permittivity varies around a stationary mean value, Λ is written:

$$\Lambda(u, \vec{r}) = \left(\frac{1}{v^2(\vec{r})} - \frac{1}{c^2} \right) \frac{\partial^2 u}{\partial t^2}$$

making it possible to rewrite the propagation equation within the Born approximation as:

$$\nabla^2 u_p - \frac{1}{c^2} \frac{\partial^2 u_p}{\partial t^2} = (\epsilon_r(\vec{r}) - 1) \frac{1}{c^2} \frac{\partial^2 u_0}{\partial t^2}. \quad [1.8]$$

Energy scattering is still present, but the rays remain unchanged:

– when permittivity varies slowly (and thus relation [1.6] is greatly verified), we will assume that the notions of wavefront and propagation ray are still valid. In this case, the solution of equation [1.4] is the geometric optical solution, which by applying Fermat's principle makes it possible to establish the curvilinear abscissa s along a ray through the relation:

$$\frac{d}{ds} \left(n \frac{d\vec{r}}{ds} \right) = \nabla n$$

Once we are positioned along the ray thus defined, our search for a solution of the type:

$$u = u(\vec{r}) e^{i\Psi(k_c \vec{r})}$$

where $k_c = 2\pi \frac{f_c}{c}$ yields non-trivial solutions if Ψ verifies the eikonal equation:

$$(\nabla \Psi(k_c \vec{r}))^2 = \epsilon(\vec{r}) \mu c^2. \quad [1.9]$$

1.1.2.2. The propagation equation in an absorbing medium

To account for a potential absorption of the incident wave, we can model the source term $\Lambda(u, \bar{r})$ using integrodifferential operators. In this case, the wave vector may formally have an imaginary component: the wave undergoes a generally significant absorption phenomenon that may even lead to a quasi-total lack of propagation (perfect conductor case).

1.1.3. Application to satellite radars

In order to reach the ground, the electromagnetic radiation emitted by the radar has to travel across the ionosphere, then the neutral atmosphere.

The ionosphere is the region of the atmosphere traditionally extending from 50 to 1,000 km in height, where there are enough free electrons to modify wave propagation. It is made up of three distinct layers, every one of which has a different electronic density ρ , expressed in electrons per cubic meter. Some of the diurnal and nocturnal characteristics of these layers are summarized in Table 1.1.

Ionosphere	Height	Diurnal ρ	Nocturnal ρ
Layer D	from 50 to 90 km	$<10^9$	~ 0
Layer E	from 90 to 150 km	$\sim 10^{11}$	$\sim 10^9$
Layer F	beyond 150 km	$\geq 10^{12}$	$\geq 10^{11}$

Table 1.1. The different layers of the ionosphere: diurnal and nocturnal rough estimates of the electronic density ρ expressed in electrons per m^3

In the ionosphere, the index depends on electronic density and is expressed for an f frequency in the form of:

$$n(f) = 1 - \frac{f_0^2}{2f^2} \quad [1.10]$$

where f_0 is the plasma frequency that depends on electronic density and can be approximated by relation $f_0 \approx 9\sqrt{\rho}$ (f_0 in MHz). Given the ionospheric ρ rough

estimates, we see that, in the different layers, this phenomenon has a very weak or even negligible effect on the centimetric waves used by imaging radars.

The neutral atmosphere, defined in terms of thermodynamic concepts, is the gaseous envelope that surrounds the Earth. It consists of several different layers: the troposphere extending from 0 to 15 km in height, where atmospheric heterogeneities are located; the stratosphere from 15 to 50 km, where temperature increases with height; and the mesosphere from 50 to 80 km¹, where temperature decreases with increasing height. Within the neutral atmosphere, the n index is essentially a function of pressure, temperature and partial water vapor pressure. Since this index is very close to 1, we will instead use refractive co-index N , which is defined as $N = 10^6(n-1)$ and expressed by a semi-empirical relation known as the Smith-Weintraub equation [LAV 97]:

$$N(T, P, e) = 77.6 \frac{P}{T} - 5.6 \frac{e}{T} + 0.375 \cdot 10^6 \frac{e}{T^2} \quad [1.11]$$

where T is the temperature in Kelvin, P is the atmospheric pressure and e is the partial water pressure², both expressed in hPa.

Systematic soundings of the entire surface of the Earth have provided an (at least statistically) accurate knowledge of the index for the stratosphere and mesosphere. In particular, Bean's atlas [BEA 66], which for most of the Earth provides the co-index $N(h)$ according to altitude h in the shape of a 5-parameter model whose coefficients are monthly averages calculated over a period of 5 years. By contrast, major index fluctuations are found in the troposphere, mostly as a result of air moisture and related clouds.

Index variations are low for both the ionosphere and neutral atmosphere: the hypotheses required by the eikonal equation are fully justified and the effects linked to index variation are only perceptible in time-of-flight measurement between the radar and the ground. In the neutral atmosphere, some gaseous components may exhibit resonances within the range of frequencies that is of interest to us, as peripheral electrons of their atoms and molecules are excited. This is the case with water vapor in particular (lines at 22.2 GHz, 183.3 GHz and 325.4 GHz) and oxygen (lines from 50 to 70 GHz and one isolated line at 118.74 GHz). The signal is almost entirely absorbed at these frequencies.

¹ The presence of free electrons in the mesosphere explains some overlapping with the ionosphere.

² This partial pressure also depends on P and T .

Other phenomena such as hydrometeors may also have an influence on propagation, both on wave delay and wave absorption. Hail, snow and lightning impact considerably on the radar signals but are difficult to model.

1.2. Matter-radiation interaction

The source term in equation [1.4] shows that the propagation of an electromagnetic wave is scattered if the medium is not homogenous. A scattered wave then appears, which is not necessarily isotropic and its radiation pattern depends on the source term. This approach does not easily cover phenomena related to the discontinuities of the propagation medium (e.g., surfaces between two media with different indices) or those related to reflecting targets.

Any phenomenological approach must take into account the radiation wavelength and L , the characteristic length of discontinuities. Even though the general case eludes all analytical considerations – except for the homogenous spherical scatterer treated by the exact Mie model – we can still analytically handle two essentially opposite cases:

- $L \gg \lambda$: this is the case of the flat interface, which can be considered as unbounded so that it possible to use the Snell-Descartes equation;
- $L \ll \lambda$: this is the case of a point target which we will refer to as a Rayleigh target.

However, this approach fails to provide a reasonable account of reality, where in rare cases there may be only one perfectly smooth surface or only one quasi-point target. A pragmatic view of reality will thus prompt us to study in more detail two cases of high practical relevance, namely rough surfaces and point target distributions.

1.2.1. Theoretical backscattering models

1.2.1.1. Unbounded flat interface

Flat interfaces have been studied since the time of Descartes and Snell. The relations obtained in visible optics (Fermat's principle) are derived from the continuity relations imposed on Maxwell equation solutions. In the case of a flat interface between two media defined by their indexes n and n' , if an incident wave impinges this interface at an angle θ with respect to the normal of the interface, we will have a reflected wave at an angle θ and, in the second medium, a wave refracted at an angle θ' , so that:

$$n \sin \theta = n' \sin \theta'$$

The only situation in which the Snell-Descartes formalism may be altered is when the second medium is more or less conductive. The wave vector will then include an imaginary component specific to attenuation through the second medium. In the case of a perfect conductor, we will only have an evanescent wave in this second medium, as the energy of the incident wave is entirely conveyed to the reflected wave.

In reality, this is obviously an ideal situation, since interfaces are neither unbounded nor rigorously flat. We may nevertheless consider that an interface can be locally put into its tangent plane: the dimensions on which this approximation is valid correspond to the dimensions of an antenna with a directivity pattern that is directly related to these dimensions. The incident wave will therefore be backscattered with some directivity pattern mainly along the refracted and reflected rays of Snell-Descartes law, as well as according to a radiation pattern for the other directions, in keeping with the Huygens principle and the diffraction theory. This radiation pattern is that of an antenna whose dimensions are those of the approximation area. Despite its simplistic appearance, this analysis gives us the order of magnitude of the backscattered field in other directions than those of the Snell-Descartes angles.

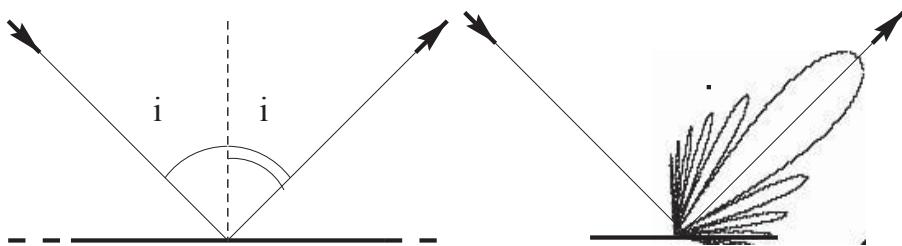


Figure 1.1. Descartes laws for an unbounded plane (left) with the same original conditions, reflection on a plane sector (right)

Note that this approach only covers the kinematic aspects. To go into finer details and include the dynamic aspects, wave polarization will need to be considered (as seen in section 1.3) which may require adding a 180° phase rotation in some cases.

1.2.1.2. Rayleigh scatterer

This case, based on a target much smaller than wavelength λ , is the opposite of the previous one. By considering a spherical homogenous target with an electric permittivity ϵ , we can resort to either exact calculations using the Mie model that makes no assumption as to sphere size, or we can choose approximate calculations (the Rayleigh model in which a sphere much smaller than the wavelength is implied).

The behavior of the Rayleigh model can be deduced from equation [1.8]. Indeed, if the target is homogenous and small enough compared to the wavelength, the source term varies little inside the target. Assuming an incident plane wave \vec{E}_i it can be written:

$$(\epsilon_r - 1) \frac{1}{c^2} \frac{\partial^2 \vec{E}_r}{\partial t^2} \sim (\epsilon_r - 1) \frac{1}{\lambda^2} \vec{E}_i,$$

where the proportionality factor involves V , i.e. the target volume.

In this way, here we have a secondary source that radiates like a dipole, proportionally to the frequency square, to local permittivity variation inside the target, and to target volume.

1.2.2. Phenomenological backscattering models

While the physical models described above provide a better understanding of how electromagnetic waves propagate, they do not make it possible to cover situations found in radar imaging. A pragmatic approach will lead us to consider three more realistic cases that will turn out to be very important for imaging: the rough interface, the case of a generic target and the scattering by a set of (point or not) targets.

1.2.2.1. Rough interface

The Snell-Descartes laws assume the interface to be flat. Such an assumption can be called into question in radar imaging as deviations from planarity must not exceed a wavelength fraction (typically $\frac{\lambda}{20}$): for example, a roughcast wall may no longer be considered, for some radar wavelengths, to be a smooth surface.

An interface is said to be rough for an incident ray impinging it at an angle θ , if the mean quadratic deviation of surface irregularities, Δh , verifies the Rayleigh quality criterion:

$$\Delta h > \frac{\lambda}{8\cos(\theta)} \quad [1.12]$$

meaning a mean quadratic phase shift higher than $\frac{\pi}{2}$.

The higher the roughness, the more the backscattering diagram differs from that of a flat interface. Moreover, it depends on the angle of incidence θ , in particular, the wider the angle of incidence, the more significant the roughness, and the more perturbed the radiation diagram.

The limit case is a surface whose roughness effects completely offset the flat interface appearance. Such a surface will then scatter incident radiation isotropically in a half-plane. The backscattering will in this case be characterized by the albedo, which represents the energetic fraction of the received signal backscattered by this surface.

1.2.2.2. Scattering by a generic target

A target that does not satisfy the Rayleigh criterion can still be characterized by using its directivity pattern and by its *Radar Cross-Section* (RCS). In order to define RCS, we will consider that the target behaves at reception like an antenna having an area a and as if the entire intercepted power was backscattered isotropically (unit gain antenna); the value of a is RCS³.

The major drawback of this model lies in the fact that RCS is often strongly dependent on the configuration under which the target is illuminated by the incident wave. Even a minor change in this configuration may cause a major change in σ .

1.2.2.3. Scattering by a set of targets

Let us consider a set of Rayleigh point targets (they can be seen as isotropic targets). These targets may be distributed on a plane (we then refer to their area density) or in a volume (we then refer to their volume density).

The backscattered wave is the sum of basic waves backscattered by every target. Assuming that target density is not too high, we will be able to omit multi-reflection,

³ This concept deserves a more elaborate definition, especially one that includes polarimetry (see, for example, [LEC 89]).

in which backscattered waves are in turn backscattered by other targets. This often justified assumption verifies the Born approximation hypotheses.

Generally, the emitted radar wave train is far longer than the wavelength: we thus speak of a coherent illumination. In this case, the sum of echoes backscattered by each target will be carried out coherently, i.e. amplitudes are summed up rather than energies⁴. The received signal therefore has a specific appearance induced by speckle, generally well known by opticians. This issue, which is a major one for radar image processing, will be discussed in more depth in Chapter 5.

1.3. Polarization

1.3.1. Definitions

When a plane divides a space into two semi-unbounded, isotropic, homogenous media, the incidence plane of an electromagnetic wave characterized by its wave vector \vec{k} can be defined as the plane containing both \vec{k} and the normal to the boundary plane dividing the two media.

The polarization of an electromagnetic wave is conventionally defined by the direction of a field \vec{E} : we say that the polarization is perpendicular if the field \vec{E} is perpendicular to the plane of incidence (TE polarization, \vec{E}_{\perp}), and that the polarization is parallel if the field \vec{E} belongs to the plane of incidence (TM polarization, \vec{E}_{\parallel}).

Starting from the Descartes laws and energy conservation, we can calculate the transmission coefficient t and reflection coefficient r for the flat interface. The fields \vec{E}_r and \vec{E}_t are related to the incident field \vec{E}_i by the following relations:

$$\begin{pmatrix} E_{r,\perp} \\ E_{r,\parallel} \end{pmatrix} = \begin{pmatrix} e^{i\vec{k}\vec{d}} \\ |\vec{d}| \end{pmatrix} \begin{pmatrix} r_{\perp} & 0 \\ 0 & r_{\parallel} \end{pmatrix} \begin{pmatrix} E_{i,\perp} \\ E_{i,\parallel} \end{pmatrix}$$

$$\begin{pmatrix} E_{t,\perp} \\ E_{t,\parallel} \end{pmatrix} = \begin{pmatrix} e^{i\vec{k}\vec{d}} \\ |\vec{d}| \end{pmatrix} \begin{pmatrix} t_{\perp} & 0 \\ 0 & t_{\parallel} \end{pmatrix} \begin{pmatrix} E_{i,\perp} \\ E_{i,\parallel} \end{pmatrix}$$

⁴ We have a very different situation where optical wavelengths are concerned, since on the one hand photons differ in frequency and, on the other hand, have very short coherence lengths. To most receivers, they will appear incoherent, which makes it possible to sum up their intensity contributions resulting in the speckle-free images we are familiar with.

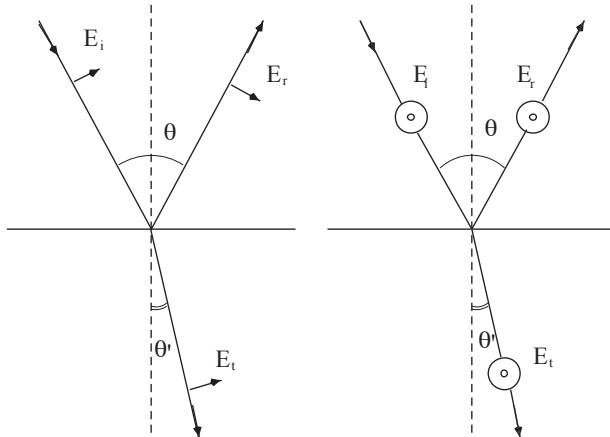


Figure 1.2. Fresnel laws for parallel polarization, i.e. where \vec{E}_i is parallel to the incidence plane (left), and for perpendicular polarization, i.e. where \vec{E}_i is perpendicular to the incidence plane (right)

where $\vec{\mathbf{d}}$ is the observer's position and [FRA 70]:

$$\begin{aligned} r_{\parallel} &= -\frac{\tan(\theta - \theta')}{\tan(\theta + \theta')}, & t_{\parallel} &= \frac{2 \cos(\theta) \sin(\theta')}{\sin(\theta + \theta') \cos(\theta + \theta')} \\ r_{\perp} &= -\frac{\sin(\theta - \theta')}{\sin(\theta + \theta')}, & t_{\perp} &= \frac{2 \cos(\theta) \sin(\theta')}{\sin(\theta + \theta')}. \end{aligned}$$

These relations highlight the different behaviors of parallel and perpendicular polarizations. In particular, for $\theta + \theta' = \frac{\pi}{2}$ the parallel polarized wave is no longer reflected; θ is known as the Brewster angle in this case.

In the general backscattering case, the components of the backscattered field \vec{E}_r , are linearly related to the incident field \vec{E}_i components. This is usually written in matrix form as follows:

$$\begin{pmatrix} E_{r,\perp} \\ E_{r,\parallel} \end{pmatrix} = \begin{pmatrix} e^{j\vec{k}\vec{\mathbf{d}}} \\ |\vec{\mathbf{d}}| \end{pmatrix} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} E_{i,\perp} \\ E_{i,\parallel} \end{pmatrix} \quad [1.13]$$

This formulation can be used to describe both the reflection by a plane and the scattering by a target, even though in the latter case parallel and perpendicular polarizations are entirely arbitrary notions.

1.3.2. Wave polarization

The polarization of a plane wave describes, versus time, the tip location of an electric field vector $\vec{E}(t)$ in a plane orthogonal to \vec{k} . Generally, this location is an ellipse (the wave is said *elliptically polarized*), which in some cases may degenerate into a straight line segment (linear polarization) or a circle (circular polarization). An elliptically polarized wave is shown in Figure 1.3 [ULA 90].

For an observer, the ellipse orientation angle ψ is the angle between the horizontal and the major axis of the ellipse describing the polarized wave. It ranges between 0° and 180° . χ is the ellipticity angle, such that, by definition the tangent is the ratio of the ellipse's minor and major axes. It ranges between -45° and $+45^\circ$, and its sign conventionally determines the direction of polarization: right if $\chi < 0$ or left if $\chi > 0$. Note that opticians refer to polarization as being positive when an observer looking at the wave that propagates towards him sees the ellipse described in the direct sense, i.e., to the left.

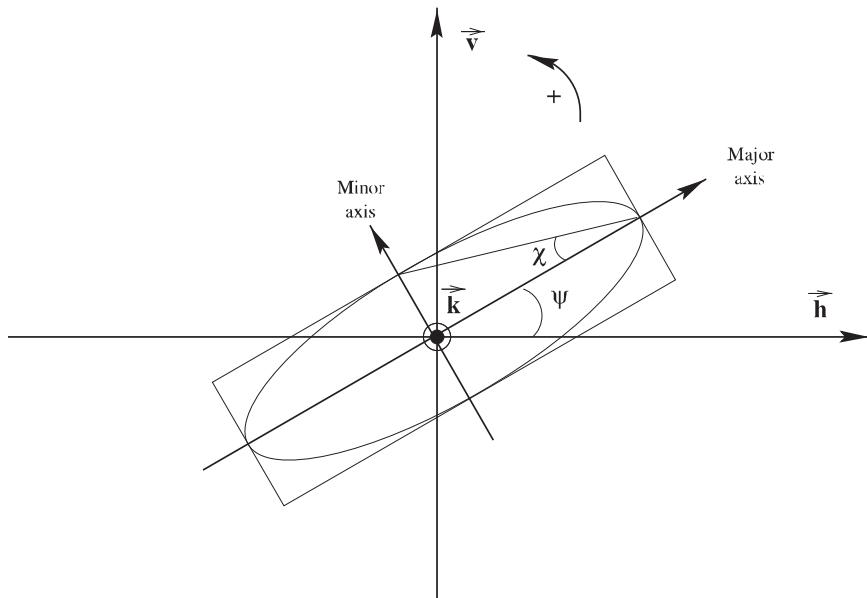


Figure 1.3. Polarization of a wave: conventions and notations

The polarization of a wave is then defined using the couple (ψ, χ) deduced from the variations of the E_h and E_v components of field $\vec{E}(t)$ along the axes h and v . These axes are defined in the plane orthogonal to \vec{k} and are conventionally related to the observer's reference frame (rather than according to its belonging to an incidence plane related to an interface, as in the previous section):

$$\begin{aligned} E_h(r,t) &= |E_h| \cdot \cos(\vec{k}\vec{r} - \omega t + \delta_h), \\ E_v(r,t) &= |E_v| \cdot \cos(\vec{k}\vec{r} - \omega t + \delta_v), \end{aligned} \quad [1.14]$$

where δ_v and δ_h are the original phase shifts. They are linked to (ψ, χ) by parameters $\Delta\delta = \delta_v - \delta_h$ and $\tan(\zeta) = |E_v|/|E_h|$, and relationships [BOR 65]:

$$\begin{aligned} \tan(2\psi) &= \tan(2\zeta) \cdot \cos(\Delta\delta) \\ \sin(2\chi) &= \sin(2\zeta) \cdot \sin(\Delta\delta) \end{aligned} \quad [1.15]$$

In the case of remote sensing radars, the observer's reference frame is related to the Earth and vector \vec{h} is horizontal. Particular cases are:

$-\chi = 0^\circ$ corresponds to linear polarizations:

$$\begin{cases} \psi = 0^\circ: \text{horizontal polarization} \\ \psi = 90^\circ: \text{vertical polarization} \end{cases}$$

$-\chi = \pm 45^\circ$ corresponds to circular polarizations:

$$\begin{cases} \chi = -45^\circ: \text{right-hand polarization} \\ \chi = +45^\circ: \text{left-hand polarization} \end{cases}$$

The polarization of a wave can also be described by using a real Stokes vector \vec{F} , defined as follows:

$$\vec{\mathbf{F}} = \begin{pmatrix} g_0 \\ g_1 \\ g_2 \\ g_3 \end{pmatrix} = \begin{pmatrix} |E_h|^2 + |E_v|^2 \\ |E_h|^2 - |E_v|^2 \\ 2 \cdot \text{Re}(E_h \cdot E_v^*) \\ 2 \cdot \text{Im}(E_h \cdot E_v^*) \end{pmatrix}$$

From equations [1.15], $\vec{\mathbf{F}}$ can be expressed versus the orientation and ellipticity angles (ψ, χ) as

$$\vec{\mathbf{F}} = g_0 \begin{pmatrix} 1 \\ \cos(2\chi) \cos(2\psi) \\ \cos(2\chi) \sin(2\psi) \\ \sin(2\chi) \end{pmatrix} \quad [1.16]$$

The first component $g_0 = |E_h|^2 + |E_v|^2$ is the total power carried by the wave. When the wave is fully polarized, i.e. the parameters $|E_h|$, $|E_v|$, δ_h and δ_v are constant over time, the wave checks the equality: $g_0^2 = g_1^2 + g_2^2 + g_3^2$ (derived from equation [1.16]). This is generally the case for the transmitted wave. Conversely, a backscattered wave is the coherent sum of waves backscattered by elementary targets (that are assumed randomly distributed) in a resolution cell, and it is represented by a random time variable. It verifies the inequality: $g_0^2 \geq g_1^2 + g_2^2 + g_3^2$, where g_i are time averages; the wave is then said to be partially polarized. The polarization degree of a wave, defined as $\sqrt{g_1^2 + g_2^2 + g_3^2} / g_0$, is therefore 1 for a completely polarized wave, less than 1 for a partially polarized wave and 0 for a completely depolarized wave.

1.3.3. The BSA convention

When an electromagnetic wave is scattered by a target, the fields are expressed in local coordinate systems related to the transmitting antenna ($\vec{\mathbf{h}_e}, \vec{\mathbf{v}_e}, \vec{\mathbf{k}_e}$) and the receiving antenna ($\vec{\mathbf{h}_r}, \vec{\mathbf{v}_r}, \vec{\mathbf{k}_r}$), while the global system is that of the observed target, as shown in Figure 1.4. In the monostatic case, i.e. when the transmission and

reception locations are the same, the variables $(\vec{h}_e, \vec{v}_e, \vec{k}_e)$ and $(\vec{h}_r, \vec{v}_r, \vec{k}_r)$ coincide according to the backscattering alignment (BSA) convention.

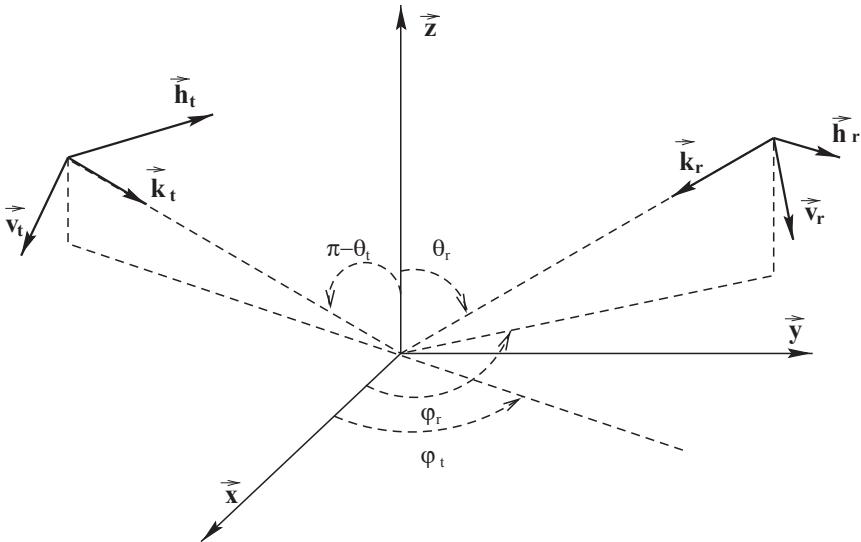


Figure 1.4. Local coordinate systems and geometry of the BSA convention, describing the incident wave and target scattered wave

1.3.4. Complex backscattering matrix S , Mueller matrix

In the case of backscattering by a target, the reflected wave and the incident wave are related to each other by the scattering matrix S (equation [1.13]):

$$\begin{pmatrix} E_{r,h} \\ E_{r,v} \end{pmatrix} = \frac{e^{i\vec{h}\cdot\vec{d}}}{|\vec{d}|} \begin{pmatrix} S_{hh} & S_{hv} \\ S_{vh} & S_{vv} \end{pmatrix} \begin{pmatrix} E_{i,h} \\ E_{i,v} \end{pmatrix} \Leftrightarrow \vec{E}_r = \frac{e^{i\vec{h}\cdot\vec{d}}}{|\vec{d}|} S \vec{E}_i \quad [1.17]$$

where \vec{d} defines the observer's location.

The elements S_{ij} of matrix S depend on the target's characteristics, particularly on the geometric (roughness) and dielectric (moisture) features, but also on acquisition characteristics, in particular wave frequency, incidence, etc. In addition, the reciprocity principle [TSA 85] implies that $S_{hv} = S_{vh}$ (rigorously this is true only when the polarized waves H and V are transmitted simultaneously, which is

actually not the case in radars alternating V and H transmissions. However, even in the latter case, data are calibrated in order to verify the relationship $S_{hv} = S_{vh}$).

In the following, we represent the complex backscattering matrix either using matrix form \mathbf{S} or vector form $\vec{\mathbf{S}}$:

$$\vec{\mathbf{S}} = \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vh} \\ S_{vv} \end{pmatrix} \quad [1.18]$$

For calibrated data, $\vec{\mathbf{S}}$ reduces to three components:

$$\vec{\mathbf{S}} = \begin{pmatrix} S_{hh} \\ S_{hv} \\ S_{vv} \end{pmatrix}$$

In numerous applications, our interest focuses on distributed or spread targets and their average properties, rather than point targets. This is the case for studies on farming crops, sea currents and iceberg drifting. For such studies, we would rather not use \mathbf{S} , but one of the two matrices given below:

– the complex Hermitian covariance matrix \mathbf{C} (monostatic case):

$$\mathbf{C} = \vec{\mathbf{S}}^* {}_t \vec{\mathbf{S}} = \begin{pmatrix} |S_{hh}|^2 & S_{hh} \cdot S_{hv}^* & S_{hh} \cdot S_{vh}^* & S_{hh} \cdot S_{vv}^* \\ S_{hv} \cdot S_{hh}^* & |S_{hv}|^2 & S_{hv} \cdot S_{vh}^* & S_{hv} \cdot S_{vv}^* \\ S_{vh} \cdot S_{hh}^* & S_{vh} \cdot S_{hv}^* & |S_{vh}|^2 & S_{vh} \cdot S_{vv}^* \\ S_{vv} \cdot S_{hh}^* & S_{vv} \cdot S_{hv}^* & S_{vv} \cdot S_{vh}^* & |S_{vv}|^2 \end{pmatrix} \quad [1.19]$$

Assuming reciprocity, \mathbf{C} reduces to a 3×3 matrix:

$$\mathbf{C} = \begin{pmatrix} |S_{hh}|^2 & S_{hh} \cdot S_{hv}^* & S_{hh} \cdot S_{vv}^* \\ S_{hv} \cdot S_{hh}^* & |S_{hv}|^2 & S_{hv} \cdot S_{vv}^* \\ S_{vv} \cdot S_{hh}^* & S_{vv} \cdot S_{hv}^* & |S_{vv}|^2 \end{pmatrix}$$

– the Stokes or Mueller matrix \mathbf{M} .

The Mueller (or Stokes) matrix has been defined such that polarimetric synthesis might be expressed using either fields $\vec{\mathbf{E}}$ or Stokes vectors $\vec{\mathbf{F}}$:

$$P \equiv |\vec{\mathbf{E}}_r \cdot \mathbf{S} \cdot \vec{\mathbf{E}}_i|^2 = \vec{\mathbf{F}}_r \cdot \mathbf{M} \cdot \vec{\mathbf{F}}_i. \quad [1.20]$$

Thus, by analogy with equation [1.17], \mathbf{M} is defined as the matrix connecting Stokes vectors with incident and reflected waves:

$$\vec{\mathbf{F}}_r = \mathbf{R} \cdot \mathbf{R}^t \cdot \mathbf{M} \cdot \vec{\mathbf{F}}_i \quad [1.21]$$

where $\vec{\mathbf{F}}_i$ is the transmitted (or incident) Stokes vector, $\vec{\mathbf{F}}_r$ is the received (or scattered) Stokes vector, and:

$$\mathbf{R} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -i & i \end{pmatrix}.$$

\mathbf{M} is a real 4×4 square matrix. In the monostatic case, it is symmetric (according to the reciprocity principle) and related to \mathbf{S} by [ULA 90]:

$$\begin{aligned}
M_{11} &= \frac{1}{4} \left(|S_{hh}|^2 + |S_{vv}|^2 + 2 \cdot |S_{hv}|^2 \right) \\
M_{12} &= \frac{1}{4} \left(|S_{hh}|^2 - |S_{vv}|^2 \right) \\
M_{13} &= \frac{1}{2} \operatorname{Re} \left(S_{hh} \cdot S_{hv}^* \right) + \frac{1}{2} \operatorname{Re} \left(S_{hv} \cdot S_{vv}^* \right) \\
M_{14} &= \frac{1}{2} \operatorname{Im} \left(S_{hh} \cdot S_{hv}^* \right) + \frac{1}{2} \operatorname{Im} \left(S_{hv} \cdot S_{vv}^* \right) \\
M_{22} &= \frac{1}{4} \left(|S_{hh}|^2 + |S_{vv}|^2 - 2 \cdot |S_{hv}|^2 \right) \\
M_{23} &= \frac{1}{2} \operatorname{Re} \left(S_{hh} \cdot S_{hv}^* \right) - \frac{1}{2} \operatorname{Re} \left(S_{hv} \cdot S_{vv}^* \right) \\
M_{24} &= \frac{1}{2} \operatorname{Im} \left(S_{hh} \cdot S_{hv}^* \right) - \frac{1}{2} \operatorname{Im} \left(S_{hv} \cdot S_{vv}^* \right) \\
M_{33} &= \frac{1}{2} |S_{hv}|^2 + \frac{1}{2} \operatorname{Re} \left(S_{hh} \cdot S_{vv}^* \right) \\
M_{34} &= \frac{1}{2} \operatorname{Im} \left(S_{hh} \cdot S_{vv}^* \right) \\
M_{44} &= \frac{1}{2} |S_{hv}|^2 - \frac{1}{2} \operatorname{Re} \left(S_{hh} \cdot S_{vv}^* \right)
\end{aligned} \tag{1.22}$$

1.3.4.1. Properties of \mathbf{M}

In the case of point targets and a monostatic radar, five relationships exist connecting \mathbf{M} terms, namely [VAN 87]:

$$\begin{aligned}
M_{11} &= M_{22} + M_{33} + M_{44} \\
M_{13} \cdot M_{23} + M_{14} \cdot M_{24} &= M_{11} \cdot M_{12} - M_{12} \cdot M_{22} \\
M_{13} \cdot M_{14} - M_{23} \cdot M_{24} &= M_{33} \cdot M_{34} - M_{34} \cdot M_{44} \\
M_{13}^2 + M_{23}^2 - M_{14}^2 + M_{24}^2 &= M_{11}^2 - M_{22}^2 \\
M_{13}^2 - M_{23}^2 - M_{14}^2 + M_{24}^2 &= M_{33}^2 - M_{44}^2
\end{aligned} \tag{1.23}$$

These relationships are a necessary and sufficient condition for a given Mueller matrix to have a single backscattering matrix associated with it. Therefore, a Mueller matrix only corresponds to an actual “physical” target when [1.23] is verified. Now, the mean Stokes parameters of the waves backscattered by an object varying in either time or space are related to the Stokes parameters of the transmitted wave by an average Mueller matrix $E[\mathbf{M}]$. However, as relations [1.23] are generally lost by averaging the \mathbf{M} matrices, there is no complex backscattered matrix corresponding to $E[\mathbf{M}]$ and only the first relation out of five in [1.23] is verified.

1.3.4.2. Other definitions of **M**

The most widespread definition of **M** is the one given above, but other definitions exist. In particular, the Mueller matrix is sometimes defined from the modified Stokes vector:

$$\begin{pmatrix} |E_h|^2 & |E_v|^2 & 2.\operatorname{Re}(E_h \cdot E_v^*) & 2.\operatorname{Im}(E_h \cdot E_v^*) \end{pmatrix}^t$$

The corresponding “modified” Mueller matrix **M_m** is then written as follows:

$$\begin{pmatrix} |S_{hh}|^2 & |S_{hv}|^2 & \operatorname{Re}(S_{hh}S_{hv}^*) & -\operatorname{Im}(S_{hh}S_{hv}^*) \\ |S_{vh}|^2 & |S_{vv}|^2 & \operatorname{Re}(S_{vh}S_{vv}^*) & -\operatorname{Im}(S_{vh}S_{vv}^*) \\ 2\operatorname{Re}(S_{hh}S_{vh}^*) & 2\operatorname{Re}(S_{hv}S_{vv}^*) & \operatorname{Re}(S_{hh}S_{vv}^* + S_{hv}S_{vh}^*) & -\operatorname{Im}(S_{hh}S_{vv}^* + S_{hv}S_{vh}^*) \\ 2\operatorname{Im}(S_{hh}S_{vh}^*) & 2\operatorname{Im}(S_{hv}S_{vv}^*) & \operatorname{Im}(S_{hh}S_{vv}^* + S_{hv}S_{vh}^*) & \operatorname{Re}(S_{hh}S_{vv}^* + S_{hv}S_{vh}^*) \end{pmatrix} [1.24]$$

There are other representations of polarimetric information. Some, such as the Poincaré sphere, the Jones representation, etc., are less widespread because they are more specific to some problems or interpretations. Some, such as the coherence matrices, will be dealt with in section 7.4.2.

This page intentionally left blank

Chapter 2

The Principles of Synthetic Aperture Radar

Formulated as early as 1891 by the American Hugo Gernsback, the radar principle (“Radio Detection and Ranging”) is based on the principles of electromagnetic propagation: an electromagnetic wave emitted by a source is backscattered by targets. The received signal, once analyzed, makes it possible to detect and locate these targets, by assuming that the propagation velocity of the wave remains fairly constant.

The first experiments in aircraft detection by radar date back to 1934, explaining why the British, Americans and Germans possessed such systems both on the ground and in the air during World War II. Nowadays, radars are well known systems with numerous applications, including detection, location, surveillance, telemetry, imaging, etc.

In this chapter, we will outline the principles of Real Aperture Radar (RAR) and Synthetic Aperture Radar (SAR)¹. Our approach is primarily based on traditional signal processing methods (matched filtering); other approaches such as the holographic approach (also called Fourier optics) yield similar results. In order to provide a better understanding of radar formalism, we will begin by analyzing the principles of surveillance and early imaging radars, paving the way for a simple overview of concepts such as antenna, ground swath and resolution. We will then continue with a discussion on the mathematical principles of SAR and finally geometric properties of images acquired by such a system.

Chapter written by Jean-Marie NICOLAS and Frédéric ADRAGNA.

¹ See [CIL 97] for a full glossary.

2.1. The principles of radar

Standard surveillance radars are typically found in airports or on ships. They make it possible to detect the presence of passive objects (called targets) by means of echoes which they send back in response to the emission of an electromagnetic pulse. They also make it possible to determine the positions of these objects (more precisely, their distance from the radar). Actually, they are not really imaging radars, the role of which would be to supply at a given moment an image of the observed scene. However, it will be useful to review here how they operate, if only to improve the understanding of imaging radars that were derived from them. We will not go into detail here except on their basic principle that has not changed in over 50 years, even though modern radars are quite sophisticated (they involve colors, moving object tracking, map overlays and so forth).

2.1.1. Description of surveillance radars

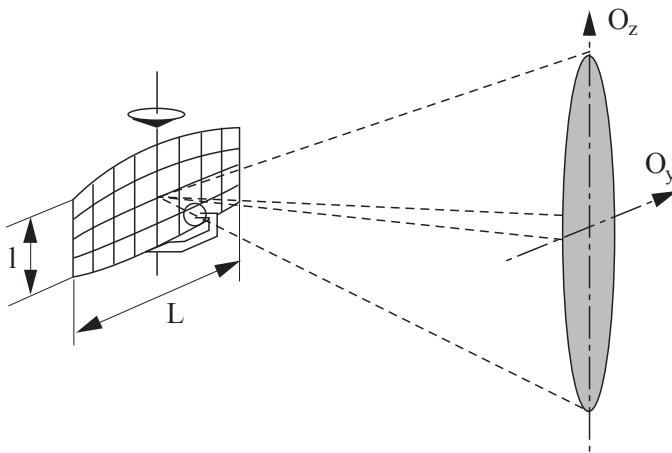


Figure 2.1. Surveillance radar: beam geometry depends on the geometric features of the antenna

A surveillance radar consists of an antenna that can rotate along the vertical axis Oz (see Figure 2.1) and a transmitter/receiver system of quasi-monochromatic microwave signals of wavelength λ . The antenna is generally long (dimension L along Oy) and not very tall (dimension l along Oz) so that $L > l$. By assimilating it to an evenly illuminated rectangular aperture, we can deduce its directivity [GOO 72].

Using the Fraunhofer approximation² that is largely valid in this case, we can write the field U observed at a point $P = \begin{pmatrix} y \\ z \end{pmatrix}$ of a plane lying at a range D from the aperture and orthogonal to \vec{D} as:

$$U(y, z) \sim \text{sinc}\left(\pi \frac{Ly}{\lambda D}\right) \text{sinc}\left(\pi \frac{lz}{\lambda D}\right)$$

The properties and shape of a sinc function are well known: the shape of the main lobe and the sidelobe locations are provided in Table 2.1 and illustrated in Figure 2.2. The sidelobes of a sinc function, with the first one at -13 dB, are always a problem in image formation, as they create artifacts on the image. In practice, however, an apodization can be performed on the radiating elements of an antenna so that the sidelobes are lowered at the expense of a broadening of the main lobe. This especially applies to slotted waveguide antennae (such as those of ERS radars and RADARSAT). In the case of active modular antennae, known as MMIC (Monolithic Microwave Integrated Circuit), energy balance constraints are such that every radiating element emits maximum energy (i.e., without apodization), even though double weighting will then be necessary on reception. Higher energy can be emitted in this way, while the antenna pattern (resulting from the transmitting and receiving patterns) remains satisfactory.

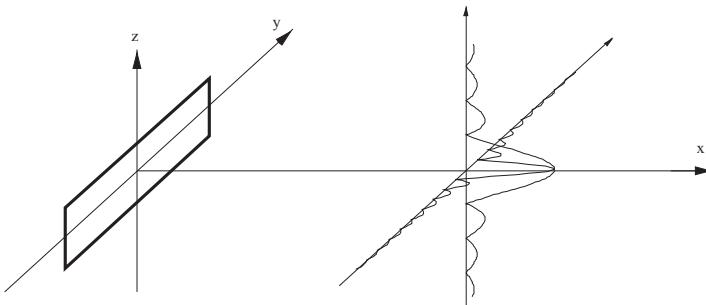


Figure 2.2. Diffraction by a rectangular antenna: energy distribution is the product of two sinc functions – a narrow one in azimuth and a wider one in range

² $D \gg \frac{L^2 + l^2}{2\lambda}$.

	$\left(\frac{ U ^2}{ U_{\max} ^2} \right)^2$	$\frac{2lz}{\lambda D}$	$\theta \sim \frac{2z}{D}$
Width at 3 dB	0.501	0.88	$0.88 \frac{\lambda}{L}$
Width at 3.92 dB	0.405	1	$\frac{\lambda}{L}$
Width at 9 dB	0.126	1.38	$1.38 \frac{\lambda}{L}$
Width between first zeros	0	2	$2 \frac{\lambda}{L}$

Table 2.1. Sinc function variations: $U(z) = \text{sinc} \left(\pi \frac{lz}{D} \right)$

The radiation pattern is also conventionally³ characterized by the angle of aperture at -3 dB : $\theta_{3 \text{ dB}} \sim \frac{\lambda}{L}$.

Even as the antenna rotates along the vertical axis, it also emits impulses in short regular succession, each time illuminating an angular space sector as narrow as possible. Following the transmission of each pulse, the radar switches into receiving mode. Any target (airplane, ship or any obstacle) located within the illuminated sector will return part of the energy it has received and the radar will detect this echo. The time (t_{AR}) it takes a wave to travel to and from the target at light speed (c) is a measure of the distance separating the antenna from the target:

$$D = \frac{ct_{AR}}{2}$$

³ This value is in fact a practical approximation: the precise variations of the sinc function are given in Table 2.1.

A radar scans the horizon (azimuth) surrounding itself and returns an image of the targets on an image scope. This image has two independent dimensions: range and azimuth, as shown in Figure 2.3.

2.1.2. Notion of resolution

A radar's resolution is linked to its ability to distinguish clearly in both range and azimuth between two adjacent targets.

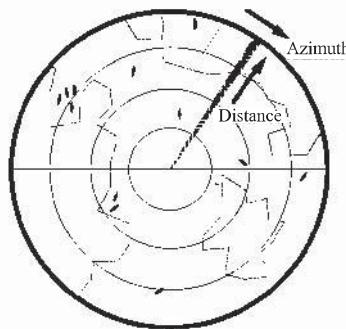


Figure 2.3. Surveillance radar: image scope

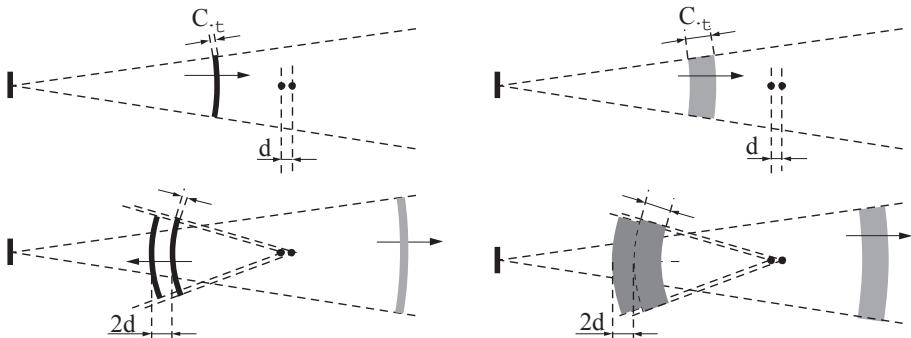


Figure 2.4. Range resolution of a radar

Range resolution r_{dist} depends on the duration of the transmitted pulse: τ (see Figure 2.4). If the pulse is very short, the radar will receive two distinct echoes from two neighboring targets (at a distance d from each other) i.e., it will be able to

distinguish between them. On the contrary, if $\tau \geq \Delta t = \frac{2d}{c}$, the echoes of the two targets will be mixed up. We will thus have:

$$r_{dist} = \frac{c\tau}{2}$$

Note that the wave travels the distance twice (to and from the target), which explains the factor 2 in these formulae.

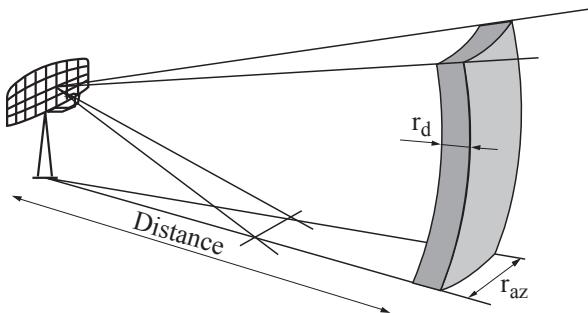


Figure 2.5. Range and azimuth resolutions of a detection radar

Azimuth resolution r_{az} (see Figure 2.5) is determined by the antenna pattern (or conventionally speaking, by the angular aperture at -3 dB) and is proportional to range:

$$r_{az} = \theta_{3 \text{ dB}} D \sim \frac{\lambda}{L} D$$

Finally, the pulse repetition frequency (PRF) is chosen so that the antenna turns by an angle $\theta = \theta_{3\text{dB}}$ between two pulse emissions.

2.1.3. Imaging radars

Unlike surveillance radars, in which image formation depends on the rotation of the system, the image obtained by imaging radar is associated with the movement of a platform bearing the side-looking antenna. Each pulse illuminates a strip of land as narrow as possible (Figure 2.6).

This is the operating principle of the first imaging radars, known as side-looking airborne radars (SLAR). This type of radar was widely used for many years, largely for military cartography purposes.

The ground swath that determines the image width depends on ranges d_1 and d_2 (Figure 2.6), which in turn depend on times t_1 , i.e. when the echo begins to be recorded, and t_2 , i.e. when the recording ends. The distance $d_1 = \frac{ct_1}{2}$ corresponding to the edge nearest to the image is called near range, whereas the distance $d_2 = \frac{ct_2}{2}$ corresponding to the opposite end of the image is called far range.

It is important to note the following points:

- As a radar deals with range information, sidelooking is required. Indeed, if we illuminated the ground vertically, we would always have two points located at the same distance, one on each side of the track (Figure 2.7). As a result, the image would fold onto itself, with points located right and left of the track mixing together.

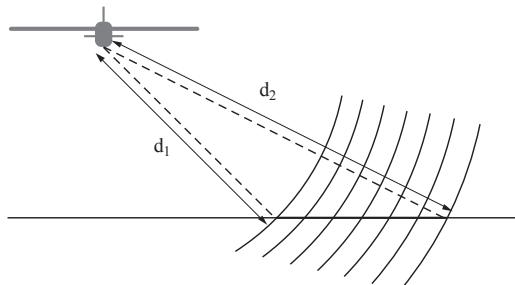


Figure 2.6. Airborne side-looking imaging radar: d_1 = near range and d_2 = far range

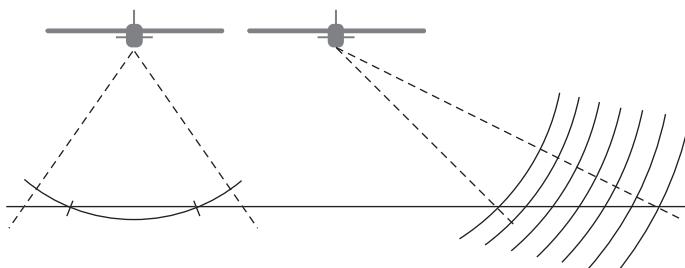


Figure 2.7. Airborne imaging radar: vertical looking and side looking.
Vertical looking leads to image folding

- Although imaging radars have a rectangular image (rather than a circular image, as provided by surveillance radars), French radar operators still use the more familiar notions of range and azimuth (see Figure 2.8). By contrast, the English-speaking radar operators would rather refer to them as *cross track* and *along track*, which is more appropriate.

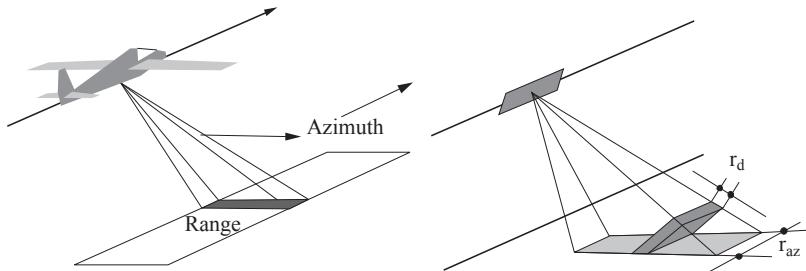


Figure 2.8. Left: range and azimuth axes of an airborne imaging radar; right: side-looking radar resolution

- Range resolution, just like above, will be proportional to the emitted pulse width $r_d = \frac{c\tau}{2}$.
- In azimuth, i.e., in the radar platform’s track direction, resolution will again depend on antenna pattern aperture and on the observation range: $r_{az} = \frac{\lambda}{L} D$. It will be rather low because an aircraft or a spacecraft cannot possibly carry a very large antenna. For example, the resolution of a SLAR with a 6 m long antenna and operating in the X-band ($\lambda = 3$ cm) will be about 50 m to 10 km range (airborne) or about 5 km to 1,000 km range (spaceborne).
- Pulse repetition frequency (PRF) has to be adapted to this resolution and to the platform speed v ; it is such that the radar travels a distance $\frac{v}{PRF} = r_{az}$ along its trajectory between two pulse transmissions.

2.1.4. Synthetic Aperture Radar (SAR)

To improve resolution, as seen in Chapter 1, the antenna has to be lengthened. Since this cannot be physically done, it takes a virtual solution to attain this goal. The American Carl Wiley first had the idea in 1951 of using platform movement and signal coherence to reconstruct a large antenna by calculation. As the radar moves

between two pulse transmissions, it is indeed possible to combine in phases all of the echoes and synthesize a very large antenna array. This is the principle of synthetic aperture radar, the mathematics of which will be discussed at length in section 2.2.3.

Such signal reconstruction depends on the precise knowledge of the platform trajectory. In the case of airborne radars, it is necessary to take into account every possible movement of the aircraft, the path of which is hardly straight and smooth, but quite influenced by yaw, roll and pitch.

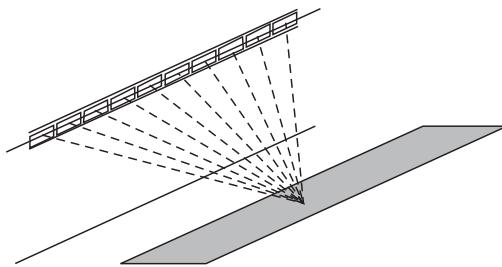


Figure 2.9. The synthetic aperture principle: all along its path, the radar acquires a series of images that are combined by post processing. The final image looks like an image acquired by an antenna that is the sum of all the basic antennae

2.1.5. The radar equation

A radar can be modeled by the so-called *radar equation* which links the received power to the transmitted power in the presence of a target characterized by RCS σ^o (see section 1.2.2.2). In the monostatic case, i.e., where the emitting antenna and the receiving antenna are the same (as they usually are in radars), we have:

$$P_r = P_e \frac{G^2 \lambda^2 \sigma^o}{(4\pi)^3 D^4 a} \quad [2.1]$$

where:

- P_r : received power
- P_e : transmitted power
- G : antenna gain
- λ : wavelength
- a : losses related to absorption in the propagation medium
- D : range between antenna and target

P	0.225–0.390 GHz	133–76.9 cm
L	0.39–1.55 GHz	76.9–19.3 cm
S	1.55–4.20 GHz	19.3–7.1 cm
C	4.20–5.75 GHz	7.1–5.2 cm
X	5.75–10.90 GHz	5.2–2.7 cm
K _u	10.90–22.0 GHz	2.7–1.36 cm
K _a	22.0–36 GHz	1.36–0.83 cm

Table 2.2. The various frequency bands, including their notations and corresponding wavelength ranges (on which, however, there is no consensus, so that considerably different values are sometimes cited in other works)

The D^4 term corresponds to geometric attenuation terms (see section 1.1.1.2) for the range traveled by the pulse from antenna to target (D^2) and back to the antenna (D^2). The RCS is a rather complex function which takes into account the dimensions (area) and the dielectric constants of the scatterer material⁴, and depends on the frequency and polarization of the incident wave.

If the radar equation shows the relation between transmitted and received wave powers, then the potential noise, whether external (radiation, Sun, etc.) or internal (circuit thermal noise, etc.), must also be taken into account as the received signals are analyzed and processed. Consequently, to assess the performance of a radar, we need to estimate the signal-to-noise ratio (SNR), which takes into consideration more factors than those mentioned earlier. These factors are generally grouped into two variables, namely, T , which represents noise input as a sole entity measured in Kelvin and the bandwidth B , and it is demonstrated that:

$$SNR \sim \frac{1}{TB}$$

This is a very important relationship and it shows that the SNR is better for a narrow bandwidth signal than for a broad bandwidth signal.

⁴ This mix of geometric and radiometric effects yield stealth properties, or non-stealth properties, of targets.

Since the Traveling Wave Tubes (TWT) commonly used on board satellites will not allow very short transmissions, separation capacity cannot be improved by acting directly on this parameter. However, *pulse compression* can work on longer pulses and still yield fine resolution by a linear frequency modulation of the signal, as will be shown in section 2.2.1.

Radar-used wavelengths cover various *bands* corresponding to the frequency ranges and wavelengths shown in Table 2.2.

In space-based remote sensing, we have to take into account a wave's capability to travel across the atmosphere. The shortest wavelengths (Ka, Ku) are strongly attenuated in the lower layers of neutral atmosphere (troposphere). Long wavelengths (P) in turn are subject to strong scattering while passing through the ionosphere (layer F). The intermediate bands (X, C, S, and L) are therefore the most widely used. Selecting one of them essentially depends on the main mission (L-band: scientific mission, biomass estimation and bio-geophysical parameters, ground penetration, polarimetry – X band: high resolution, cartography, detection) and technological limitations, including the radar's size and antenna length (which are proportional to wavelength) active module efficiency, etc. The C-band (ERS, RADARSAT) offers an acceptable tradeoff for all applications. Future missions will most likely tend to use two-frequency systems, with the two bands being as far apart as possible (X+L).

2.2. The SAR equations

We present in this section the principle of pulse compression related to range resolution, and the principle of SAR synthesis, which is linked to azimuth resolution. For the sake of clarity, simplified assumptions have been made. The obtained results (such as the relations providing the radial and azimuth resolutions) should therefore be viewed as values indicating the overall dimension of the problem and not exact results.

2.2.1. *The principle of pulse compression*

In order to better understand the principles of SAR, let us first consider a non-moving radar at a point P that emits an almost sinusoidal, frequency-modulated signal (such signals are widely known as “chirps”), centered on frequency f_c and represented by the complex quantity $A(t)$:

$$A(t) = A_0 e^{2j\pi \left(f_c t + \frac{Kt^2}{2} \right)} \quad |t| \leq \frac{\tau}{2}. \quad [2.2]$$

We assume that at range D_c there is a target backscattering with RCS σ^o . We then receive on the radar a signal $v_r(t)$:

$$v_r(t) = \sigma^o A_0 e^{2j\pi \left(f_c(t-t_c) + \frac{K(t-t_c)^2}{2} \right)} \quad |t-t_c| \leq \frac{\tau}{2}$$

where $t_c = \frac{2D_c}{c}$. To be able to analyze the received signal, we filter it using a filter matched to the emitted signal $A(t)$. The result is a detected signal $g_{t_c}(t)$:

$$\begin{aligned} g_{t_c}(t) &= \int_{-\infty}^{+\infty} A^*(t'-t) v_r(t') dt', \\ &= \sigma^o A_0^2 e^{2j\pi f_c(t-t_c)} \int_{-\infty}^{+\infty} e^{-2j\pi \frac{K(t'-t)^2}{2}} e^{2j\pi \frac{K(t'-t_c)^2}{2}} dt'. \end{aligned}$$

where t' verifies the double constraint that:

$$(t' \Leftrightarrow t) \in \left[-\frac{\tau}{2}, \frac{\tau}{2} \right] \quad \text{and} \quad (t' \Leftrightarrow t_c) \in \left[-\frac{\tau}{2}, \frac{\tau}{2} \right].$$

We address here the case where $t > t_c$ (the other case is written analogously) and have $t' \in \left[t - \frac{\tau}{2}, t_c + \frac{\tau}{2} \right]$, so we can write:

$$g_{t_c}(t) = \sigma^o A_0^2 e^{2j\pi f_c(t-t_c)} \int_{t-\frac{\tau}{2}}^{t_c+\frac{\tau}{2}} e^{-2j\pi K \frac{(t'-t)^2}{2}} e^{2j\pi K \frac{(t'-t_c)^2}{2}} dt'$$

Considering that $t_2 = \frac{t+t_c}{2}$, $T_2(t) = \frac{\tau}{2} - \frac{t-t_c}{2}$ (hence verifying $T_2(t) \leq \frac{\tau}{2}$) and $u = 2\pi K(t - t_c)$, we obtain:

$$g_{t_c}(t) = \sigma^o A_0^2 e^{2j\pi f_c(t-t_c)} 2T_2(t) \frac{\sin(uT_2(t))}{uT_2(t)}.$$

The general expression of the detected signal, i.e., for $t \in [t_c - \tau, t_c + \tau]$ is written:

$$g_{t_c}(t) = \sigma^o A_0^2 e^{2j\pi f_c(t-t_c)} 2T_s(t) \frac{\sin(uT_s(t))}{uT_s(t)}, \quad [2.3]$$

$$\text{where } T_s(t) = \frac{\tau}{2} - \frac{|t-t_c|}{2}.$$

The result is a signal whose shape is a sinc function weighted by term $T_s(t)$, representing a triangular window 2τ in width and centered on t_c . This window is particularly important in the case where the wave is not frequency modulated (then $K = 0$ and the sinc function is reduced to a unit pulse).

The argument of the sinc function depends on t by u and $T_s(t)$. In fact, it is verified that for the common values of K and r in an area V_{t_c} around t_c , we can only take u into account, since T_s varies very slowly and can be reasonably well approximated by the constant value $\frac{\tau}{2}$. In this case equation [2.3] becomes:

$$\begin{aligned} g_{t_c}(t) &\approx \sigma^o A_0^2 e^{2j\pi f_c(t-t_c)} \tau \frac{\sin(u \frac{\tau}{2})}{u \frac{\tau}{2}} \\ &\sim e^{2j\pi f_c(t-t_c)} \frac{\sin(\pi K \tau (t-t_c))}{\pi K \tau (t-t_c)}, \quad \text{for } t \in V_{t_c}. \end{aligned} \quad [2.4]$$

2.2.2. Range resolution

Resolution, i.e., the ability to distinguish between two targets, is obtained by analyzing the sinc function. It is characterized by $\delta t = t - t_c$, which corresponds to range $\frac{c \delta t}{2}$ (where the factor $\frac{1}{2}$ is introduced to cope with the round-trip propagation). While the concept was treated in a simplified approach in section 2.1.2, we can in fact use several criteria to define how two targets are distinguished. These criteria lead to different resolutions presented in Table 2.3 according to the selected criterion.

	$\left(\frac{g_{t_c}(t)}{g_{\max}} \right)^2$	$u \frac{\tau}{2} = K \tau \delta t \pi$	δt chirp	δt without chirp
Width at 3 dB	0.501	0.88π	$\frac{0.88}{K \tau}$	0.58τ
Width at 3.92 dB	0.405	π	$\frac{1}{K \tau}$	0.73τ
Width at 9 dB	0.126	1.38π	$\frac{1.38}{K \tau}$	1.28τ
Width between zeros	0	2π	$\frac{2}{K \tau}$	2τ

Table 2.3. Range resolution: for each space resolution criterion, the necessary time resolution is provided either for the case of a frequency modulated wave (chirp) or for the case of a purely sinusoidal wave limited by the triangular window $T_s(t)$ alone

We also indicate the width of the triangular window that would lead to this very resolution if the transmitted signal were pure frequency. As can be seen, for equal pulse widths, the pulse compression principle – here applied by chirp – improves our resolution by a factor in the order of $\frac{1}{K \tau^2}$.

We will use the lobe width of 3.92 dB as a reference hereafter, which will result in a range resolution:

$$\delta_d = \frac{c}{2} \frac{1}{K\tau} = \frac{c\tau}{2} \frac{1}{K\tau^2}$$

It is important to note that range resolution is independent from frequency f_c and depends solely on the bandwidth $K\tau$ of the transmitted signal.

In order to better analyze the resolution characteristics of a SAR, we substitute the variable $x = \frac{ct}{2}$ in equation [2.2]. We then have:

$$A(t) = A_0 e^{2j\pi \left(\frac{2x}{\lambda} + \frac{\tilde{K}_d x^2}{2} \right)} |x| \leq \frac{\tilde{X}_d}{2}$$

where $\tilde{K}_d = \frac{4K}{c^2}$ and $\tilde{X}_d = \frac{c\tau}{2}$.

For a target with $x_c = \frac{c\tau}{2}$, equation [2.4] is written as:

$$\tilde{g}_{x_c}(t) \sim e^{2j\pi \frac{2(x-x_c)}{c}} \frac{\sin(\pi \tilde{K}_d \tilde{X}_d (x \Leftrightarrow x_c))}{\pi \tilde{K}_d \tilde{X}_d (x \Leftrightarrow x_c)}, \quad \text{for } x \in \nu_{x_c}.$$

Application to the ERS-1 satellite

In the case of ERS-1, we have the following numerical values.

K	$0.41889 \cdot 10^{12} \text{ s}^{-2}$	
R	$37.12 \cdot 10^{-6} \text{ s}$	chirp duration
$K\tau$	15.55 MHz	
$K\tau^2$	577.22	compression factor
$\frac{1}{K\tau}$	64.310^{-9} s	
$\delta_d = \frac{c}{2} \frac{1}{K\tau}$	9.65 m	range resolution

Using a chirp leads to a gain in the order of $K\tau^2 = 577$, meaning that resolution would have been of several kilometers without pulse compression.

In the radar geometry reference frame (distance reference frame, where $x = \frac{ct}{2}$), the following values are obtained:

\bar{K}_d	$1.862 \cdot 10^{-5} \text{ m}^{-2}$	
\bar{X}_d	$5.568 \cdot 10^3 \text{ m}$	
$\bar{K}_d \bar{X}_d$	10.37 m^{-1}	
$\bar{K}_d \bar{X}_d^2$	577.22	compression factor
$\delta_d = \frac{c}{2} \frac{1}{K\tau}$	9.65 m	range resolution

2.2.3. Azimuth resolution

2.2.3.1. Principles

Returning to equation [2.4] obtained for a target located at range D_c :

$$g_{t_c}(t) \sim e^{2j\pi f_c(t-t_c)} \frac{\sin(u\frac{\tau}{2})}{u\frac{\tau}{2}}, \quad \text{for } t \in \nu_{t_c}$$

Since $t_c = \frac{2D_c}{c}$, D_c variations will affect $g_{t_c}(t)$ in two different ways:

– in the exponential term: $e^{2j\pi f_c(t-t_c)}$. The higher f_c is, the faster the phase variation. Particularly, any D_c variation equal to λ will result in a phase rotation by 2π .

– in the term corresponding to the sinc function: $\frac{\sin(u\frac{\tau}{2})}{u\frac{\tau}{2}}$. In fact, the shape of this function has virtually no variation and its peak is obtained for $u = 0$, i.e., at a time $t = t_c = \frac{2D_c}{c}$.

Therefore, we notice that the prevailing term is $\tilde{g}_{t_c}(t)$:

$$\tilde{g}_{t_c}(t) \sim e^{2j\pi f_c(t-t_c)} = e^{2j\pi f_c\left(t-\frac{2D_c}{c}\right)}$$

Now consider basis (Oxy) , where axis Oy is related to the platform track, and with the target on axis Ox , which is orthogonal to Oy and corresponds to the radar axis (the basis is the same as in Figure 2.2, where the antenna is at the origin). This basis illustrates the ideal case of a radar looking perpendicularly to its track: there is no antenna squint in this case. For a given position y of the radar within this reference, distance D_c between the radar and the target is written:

$$D_c = \sqrt{D_0^2 + y^2} \approx D_0 + \frac{y^2}{2D_0} \quad [2.5]$$

where D_0 is the distance from target to axis Oy , enabling us to write:

$$\tilde{g}_{t_c}(t) = \tilde{G}(t, y) = e^{2j\pi f_c \left(t - \frac{2D_c}{c}\right)} = e^{2j\pi f_c \left(t - \frac{2D_0}{c}\right)} e^{2j\pi \frac{f_c y^2}{D_0 c}} \quad [2.6]$$

We find that \tilde{G} in this approximation is a function that can be separated from t and y . Particularly, the term containing $y: e^{2j\pi \frac{f_c y^2}{D_0 c}}$, of phase $\phi_{az} = 2\pi \frac{f_c y^2}{D_0 c}$ can be viewed as a *chirp* centered on a (space) frequency 0 and having a constant $\tilde{K}_a = \frac{2f_c}{D_0 c} = \frac{2}{D_0 \lambda}$:

$$\tilde{g}_{t_c}(t) = e^{2j\pi f_c \left(t - \frac{2D_0}{c}\right)} e^{2j\pi \frac{\tilde{K}_a y^2}{2}}$$

The resolution obtained with matched filter can be deduced straight from Table 2.3 and entered in Table 2.4.

In Table 2.4, value \tilde{Y} is the size of an antenna whose resolution is defined by its lobe width (in dB), which provides the angular visibility of a target according to this criterion (in dB). If we choose to have the same value for the resolution of a synthetic antenna (sinc function obtained by matched filtering) as for the visibility of a target by the basic antenna (scattering laws), we can immediately deduce that:

$$\delta_a = \frac{1}{\tilde{K}_a \tilde{Y}} = \frac{L}{2} \quad [2.7]$$

What may seem striking in equation [2.7] is that azimuth resolution only depends on the antenna size, i.e. the smaller the antenna, the finer the azimuth resolution. This is essentially because a smaller antenna has a wider lobe, and a target will be seen for a longer time, which amounts to synthesizing a larger antenna, hence with a finer resolution. Let us note that we are not talking about a strict equality here, but rather about a good approximation that can be used for rapid sizing. The precise resolution will depend on the lobe width that is actually synthesized by the SAR processing and on how much weight is applied to the signal. This formula will be provided in section 2.2.3.3.

	$\left(\frac{\tilde{G}(t)}{\tilde{G}_{\max}} \right)^2$	$\tilde{K}_a \tilde{Y} \delta y \pi$	δy	\tilde{Y}	δy
Width at 3dB	0.501	0.88π	$\frac{0.88}{\tilde{K}_a \tilde{Y}}$	$0.88 \frac{D_0 \lambda}{L}$	$\frac{L}{2}$
Width at 3.92 dB	0.405	π	$\frac{1}{\tilde{K}_a \tilde{Y}}$	$\frac{D_0 \lambda}{L}$	$\frac{L}{2}$
Width at 9 dB	0.126	1.38π	$\frac{1.38}{\tilde{K}_a \tilde{Y}}$	$1.38 \frac{D_0 \lambda}{L}$	$\frac{L}{2}$
Width between first zeros	0	2π	$\frac{2}{\tilde{K}_a \tilde{Y}}$	$2 \frac{D_0 \lambda}{L}$	$\frac{L}{2}$

Table 2.4. Azimuth resolution: the integration bound \tilde{Y} is provided for an antenna of length L and an antenna sidelobe given in dB in the first column. The value in the last column δy is obtained by taking into account that $\tilde{K}_a = \frac{2}{D_0 \lambda}$

2.2.3.2. Application to the ERS-1 satellite

For ERS-1 we have the following values:

L	10 m	
D_0	880 km	
\tilde{K}_a	$4.015 \cdot 10^{-5} \text{ m}^{-2}$	
$\tilde{L} = \tilde{Y}$	4.98 km	length of synthetic antenna
$\tilde{K}_a \tilde{Y}^2$	996	compression factor
$\tilde{K}_a \tilde{Y}$	0.2 m^{-1}	
$\delta_a = \frac{1}{\tilde{K}_a \tilde{Y}}$	5 m	azimuth resolution

These values show that, by processing the signal over a 4.98 km distance, we synthesize an antenna having a length $\tilde{L} = 4.98$ km. The result at 3.92 dB is a 5 m resolution (i.e., a gain of 996 over the original 10 m antenna with a resolution of 5 km).

2.2.3.3. Taking platform motions into account

Assuming the platform is in uniform linear motion at a velocity V_s along Oy , we can write $y = V_s t$; in this case, the y^2 term in $\tilde{G}(t, y)$ can be rewritten as:

$$G_y(t) = e^{2j\pi \frac{f_c y^2}{D_0 c}} = e^{2j\pi \frac{f_c V_s^2 t^2}{D_0 c}} = e^{2j\pi \Phi(t)}$$

where the instantaneous frequency is written:

$$F = \frac{1}{2\pi} \frac{d\Phi}{dt} = \frac{2f_c V_s^2 t}{D_0 c} = \frac{2V_s^2}{\lambda D_0} t$$

$\frac{V_s t}{D_0}$ is actually a misalignment equal to $\tan \psi \approx \sin \psi$ (small angle approximation), which makes it possible to write:

$$F = f_c \frac{2V_s \sin \Psi}{c} = F_d;$$

the instantaneous frequency F being identical to the Doppler frequency F_d – a frequency offset caused by the transmitter's motion relative to the target. The most general form of Doppler frequency at a point detected by the vector \vec{D} related to the sensor is:

$$F_d = \frac{\vec{V}_s \cdot \vec{D}}{\lambda |\vec{D}|} \quad [2.8]$$

This formal analogy justifies the existence of other approaches to the synthetic aperture principle in which antenna laws are largely formulated based on the Doppler frequency of the targets.

The case of spaceborne radars can be treated as the airborne radar case, only instead of V_s we use the velocity of the point illuminated on the ground (ground antenna footprint velocity), which can be assimilated to the velocity of the satellite ground track as calculated in an Earth-related frame (fixed Earth).

In this time reference frame, we have $t \in \left[-\frac{\tau_a}{2}, \frac{\tau_a}{2} \right]$, where $\tau_a = \frac{\tilde{Y}}{V_s}$. By setting $K_a = \frac{f_c V_s^2}{D_0 c} = \tilde{K} V_s^2$ we can write:

$$G_y(t) = e^{2j\pi \frac{K_a t^2}{2}}$$

We then obtain for ERS-1 the following (non-weighted) values:

K_a	2.238 10^3 s $^{-2}$	
τ_a	0.667 s	collection time
$K_a \tau_a$	1,493 Hz	Doppler band
$K_a \tau_a^2$	996	
$\frac{1}{K_a \tau_a}$	0.669 10^{-3} s	
$\delta_a = \frac{V_s}{K_a \tau_a}$	5 m	azimuth resolution

2.2.4. Transfer function of the SAR antenna

In the frame of approximations leading to the separation of variables, the transfer function of the SAR antenna can be deduced directly from the above relations; so for a target located in (x_c, y_c) it will be written:

$$h(x, y) = \frac{\sin \pi \tilde{K}_d \tilde{X}(x - x_c)}{\pi \tilde{K}_d \tilde{X}(x - x_c)} \frac{\sin \pi \tilde{K}_a \tilde{Y}(y - y_c)}{\pi \tilde{K}_a \tilde{Y}(y - y_c)} \quad [2.9]$$

2.2.5. Range and azimuth samplings

In our earlier reasoning we were assuming that the functions $g_{t_c}(t)$ (range resolution) and $\tilde{G}(t, y)$ (azimuth resolution) were continuous in t and y . In practice, the received time signal is sampled temporally and the space positions accessible by the satellite along direction Oy correspond to a discrete number of points depending on the pulse repetition frequency (*PRF*).

In carrying out both discretizations, the basic sampling principles, particularly the Shannon criterion, have to be observed. Failure to satisfy this criterion can lead to artifacts caused by spectrum aliasing.

2.2.5.1. Range sampling

The sampling rate must be high enough to provide an unambiguous representation of the chirp with frequency ranges between $f_c - \frac{K\tau}{2}$ and $f_c + \frac{K\tau}{2}$: so its bandwidth is $K\tau$. Consequently, to satisfy Shannon criterion, the sampling rate F_e needs to be higher than $K\tau$. For example, the sampling rate of the ERS 1 & 2 digitizer is $F_e = 18.96$ MHz, higher than value $K\tau = 15.55$ MHz.

2.2.5.2. Azimuth sampling

As the radar moves along its orbit, it takes a number of positions P_i on the axis Oy corresponding to moments of transmission and reception of the radar wave. These points P_i are defined by their coordinates on the axis Oy : y_i . We will assume in the following that:

- there are $2N + 1$ positions and $i \in [-N; N]$;
- P_0 corresponds to the origin of benchmark Oxy ;
- $y_i = i L_{PRF} = i V_s T_{PRF}$: so the positions are regularly sampled and $L_{PRF} = V_s T_{PRF}$ is the azimuth sampling step.

The spatial sampling rate must in turn ensure an unambiguous representation of equation [2.6] the bandwidth of which is $\frac{1}{\tilde{K}\tilde{Y}}$. To this end, it is enough that the azimuth sampling step L_{PRF} verifies $\frac{1}{L_{PRF}} \geq \tilde{K}\tilde{Y} = \frac{2}{L}$, which comes down to saying that the distance traveled by the antenna between two successive acquisitions must not exceed its theoretical resolution.

In other words, in an approach using the Doppler shift to satisfy the Shannon criterion, we may say that the desired band $\frac{1}{T_{PRF}} = PRF$ has to be larger than the Doppler band $K_a \tau_a$.

2.2.5.3. *Choice of a PRF: theoretical constraints*

As we have seen, PRF is a parameter with a value that is dictated by several considerations unrelated to one another, even clashing:

- swath size, since the useful signal is analyzed during reception, i.e., between two transmissions. In satellite systems as a rule, we want to maximize the swath value to obtain a full coverage of the Earth's surface in as few days as possible and therefore we will want to minimize PRF;
- desired azimuth resolution, since the movement of the antenna between two acquisitions must not exceed theoretical resolution, which amounts to maximizing PRF;
- monostatic use of the antenna: it is imperative in this case that the desired signal strike the antenna between two transmissions.

As will be shown in the next section, this list is unfortunately incomplete (nadir echoes will have to be taken into account), which will make the value of this parameter yet harder to determine.

2.2.6. *Operation of a satellite SAR*

The results that we demonstrated based on the radial (pulse compression) and azimuth (synthetic antenna principle) signal analysis shown in the previous section are of fundamental, yet mostly theoretical importance. Indeed, what we considered a perfect antenna in both azimuth and range directions, and our demonstrations that helped establish the antenna laws of a SAR system actually took into account only its main lobe, while the sidelobes were neglected.

In order to operate a SAR, we first have to check the validity of this hypothesis in more detail. We will therefore focus in this section on the effects of sidelobes in a real antenna and their consequences on the images obtained.

2.2.6.1. Effects on radial quality

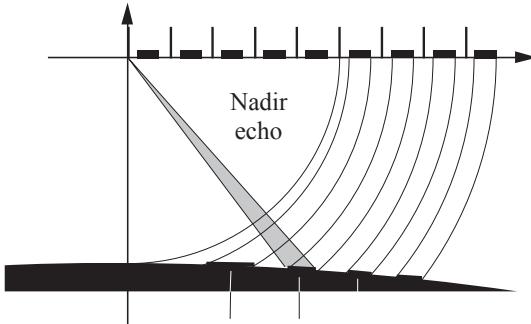


Figure 2.10. SAR image acquisition: the swath (in gray) lies within the main lobe of the antenna

The swath of an imaging radar depends, as seen in section 2.1.3, on the near range $\left(\frac{ct_1}{2}\right)$ and the far range $\left(\frac{ct_2}{2}\right)$: these values have to be determined in such a way that the corresponding points on the ground fall within the main antenna lobe. Given the distances involved in spaceborne systems and PRF values, a radar emits several pulses by the time the echo of the first one has returned. We then need to define t'_1 and t'_2 , i.e. time intervals passed since the last pulse, such that $(kT_{PFR} + t'_1) = t_1$ and $(kT_{PFR} + t'_2) = t_2$ (where $k = 9$ for ERS). The area that we are seeking to observe is thus located between distances $\frac{c(k.T_{PFR})+t'_1}{2}$ and $\frac{c(k.T_{PFR})+t'_2}{2}$. As shown in Figure 2.10, the received echoes may come from areas verifying the equation for a different k . Such echoes are called *range ambiguities* and the ambiguous areas returning them are those illuminated by antenna sidelobes. We therefore have to see, by an elevated antenna apodization for example, that they remain as faint as possible. However, this is not enough to suppress the echoes coming from the nadir⁵. As the surface of the ground is often horizontal, the waves sent in that direction return towards the radar following a specular, or very strong, reflection. Moreover, all points within a broad area in the nadir direction are equally distant from the radar and send back echoes within the same time window.

⁵ The nadir is the vertical direction opposed to zenith (hence pointing to the center of the Earth).

In order to avert such artifacts, the PRF value is chosen such that the nadir echo arrives while the antenna is in transmitting mode. This imposes yet another constraint on the way *PRF* is chosen:

$$\frac{2H}{cT_{PFR}} = p \quad p \in \mathbb{N}$$

2.2.6.2. Effects on azimuth quality

We have seen that to satisfy Shannon criterion, we need to have an azimuth phase rotation with a value lower than π between two elements of the synthetic antenna. To this end, we analyzed a point located at $y = 0$, since we assumed that signals were only returned by points lying within the main lobe of the real antenna along Oy .

In reality, this direction is not free from sidelobes either. Let us consider a point located in $y = 2Y_{max}$ (first sidelobe peak): we note that the azimuth phase difference between two synthetic antenna points ranges between π and 3π . Since the excursion of this value does not exceed 2π , the Shannon criterion is satisfied and this specific point will be reconstructed on an energy level corresponding to the value of the antenna lobe in this direction and in the same way as a point located on the antenna axis would be. Now, in the specific case where we have a weak signal in the direction of the antenna (e.g., the sea), while the signal in this particular direction is strong (e.g., a harbor), the point corresponding to the first sidelobe will be more visible than the desired point.

The network lobes are responsible for this occasionally quite annoying phenomenon which is often referred to as the Doppler frequency ambiguity due to the analogy between azimuth frequency and the Doppler shift, as seen in section 2.2.3.3.

2.2.6.3. The squint case

The calculation performed in section 2.2.3 enabled us to determine azimuth resolution in the particular case of an antenna directed perpendicularly in the trajectory of the platform. We generally notice in this case a squint characterized by an angle ψ : the lobe of the antenna is then pointed at a target with an abscissa y_c which is given by relation $y_c = D_0 \sin(\psi) \approx D_0\psi$. Approximation [2.5] has to be replaced by another one as follows:

$$D \approx D_0 + \frac{(y - y_c)^2}{2D_0}.$$

With this new approximation, the calculation proceeds exactly as described earlier. The y^2 term in $\tilde{G}(t, y)$ (equation [2.6]) can be rewritten:

$$G_y(t) = e^{2j\pi \frac{f_c(y-y_c)^2}{D_0c}}$$

and the azimuth resolution formula is obtained exactly as if there were no squint at all.

Note that for $y = 0$, the Doppler frequency will no longer be zero, but instead equal to:

$$F_{DC} = -\frac{2f_c V_s y_c}{D_0 c} = -\frac{2V_s y_c}{\lambda D_0} = -\frac{2V_s}{\lambda} \sin \Psi.$$

This frequency F_{DC} is known as Doppler centroid and varies slightly along the swath. If we do not know the exact antenna squint, the radar will not be able to provide an unambiguous image unless $|F_{DC}| \leq \frac{PRF}{2}$. This is the case with ERS (yaw steering of the platform), but not with RADARSAT.

2.3. Acquisition geometry of SAR images

The geometric characteristics of images acquired by side-looking radars (whether SAR or RAR) often appear puzzling to anyone accustomed to optic images. Indeed, the site image pixel positions are defined in terms of range: the presence of relief will cause complex distortions.

Therefore, this section will tackle the acquisition geometry of SAR systems and related distortions.

2.3.1. Sampling and resolution

As shown above, in normal imaging conditions (i.e., little back or forward squint, if any), radar imaging geometry can be broken down into two completely separable dimensions:

- range (perpendicular to the track: cross track); and
- azimuth (parallel to the track: along track).

2.3.1.1. Range

Since the propagation velocity of a radar wave can be considered constant and equal to c in a first approximation, we are able to relate times and distances by analyzing the temporal signal received between time t_1 (corresponding to the near range) and time t_2 (corresponding to the far range). If we denote by θ the mean angle of look angle incidence between the electromagnetic wave and the ground, we will obtain the following relations:

$$\begin{aligned} \text{– Near range} &= \frac{ct_1}{2} \\ \text{– Far range} &= \frac{ct_2}{2} \\ \text{– Swath} &= \frac{c(t_2-t_1)}{2\sin(\theta)} \end{aligned}$$

The sampling frequency F_e cuts the signal into samples called *gates* of duration $\frac{1}{F_e}$ and size (in radial range):

$$\Delta_d = \frac{\frac{c}{2}}{F_e}.$$

This is illustrated in Figure 2.11. Range resolution is identically expressed, but in terms of the transmitted chirp parameters:

$$\delta_d = \frac{\frac{c}{2}}{K\tau}.$$

Δ_d being generally chosen by taking into account over-sampling: $\Delta_d < \delta_d$

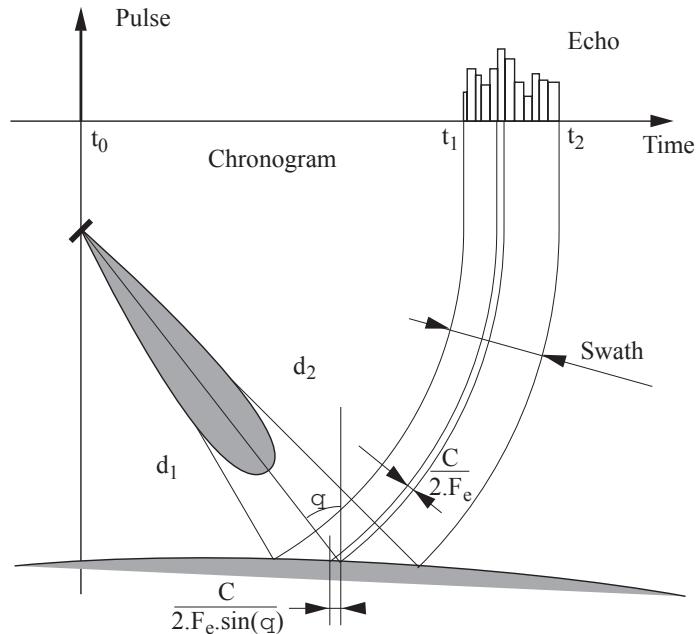


Figure 2.11. Image construction by segmenting the return signal into temporal gates

In the case of ERS-1, the values are as follows:

$K\tau$	15.55 MHz	δ_d	9.65 m
F_e	18.9627 MHz	Δ_d	7.9048 m

2.3.1.2. Azimuth

The PRF plays the part of sampling frequency in this direction. For a smooth linear trajectory, the size of an azimuth gate will be:

$$\Delta_a = L_{PFR} = V_s T_{PFR} = \frac{V_s}{PRF}$$

where V_s is the platform velocity. In the case of spaceborne radars, the orbit (which is quasi-circular) and the Earth's rotation must be taken into account. The formula remains the same if we use satellite ground track instead of platform velocity, which has its own velocity due to the Earth's rotation (relative velocity).

Azimuth resolution only depends on characteristics related to the synthetic antenna operation. It is therefore written as:

$$\delta_a = \frac{1}{L} = \frac{V_s}{K_a \tau_a}$$

In order to avert loss of information, we will pick a sampling: $\Delta_a < \delta_a$.

For a given PRF, we have the following values in the case of ERS-1:

δ_a	5 m
Δ_a	4 m

2.3.2. Radar geometry and ground geometry

2.3.2.1. Radar geometry

Gate sizes and range resolutions described in the previous section are said to be in radar geometry. They are expressed in radial range or slant range with respect to the satellite position.

2.3.2.2. Ground geometry

It is often more interesting to describe phenomena within a ground-related frame by performing a ground projection, with the ground being assumed as flat. In terms of range, gate size and resolution then depend on the wave look angle θ :

$$\Delta_{d, \text{flat ground}} = \frac{c}{2F_e \sin(\theta)}.$$

$$\Delta_{d, \text{flat ground}} = \frac{c}{2K\tau \sin(\theta)}.$$

Ground projection causes no distortion in terms of azimuth. The conventional algorithms of radar synthesis focus around the Doppler centroid that looks like a cone with its tip in the radar and the orbit as axis (according to equation [2.8]). The intersection of this cone with the ground is a hyperbola. Current algorithms synthesize raw radar data at zero frequency (zero Doppler) that corresponds to the plane perpendicular to the orbit. Radar geometry and ground geometry are then identical.

For ERS satellites, the gate size varies from 26 m at near range to 18 m at far range, while their size in azimuth remains constant at about 4 m.

2.3.3. Overlays, foreshortening and shadow

On the range axis, relief, like incidence, alters the size of gates projected on the ground (see Figure 2.12). The gate size is obtained from:

$$\Delta_{d,\alpha} = \frac{c}{2F_e \sin(\theta - \alpha)}$$

where α is the local ground slope.

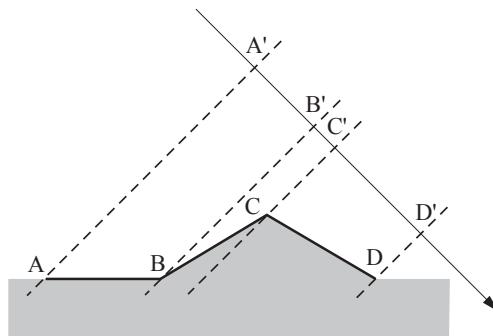


Figure 2.12. Ground slope effect on the size of gates projected on the ground

In the case of a positive slope directed towards the radar, the pixels are larger than pixels on flat ground, thus leading to a foreshortening of the image in radar geometry (see Figure 2.13). If the ground slope is equal to incidence, the range gate grows so large that the entire slope lies within one gate. When the slope is higher than incidence there is overlay.

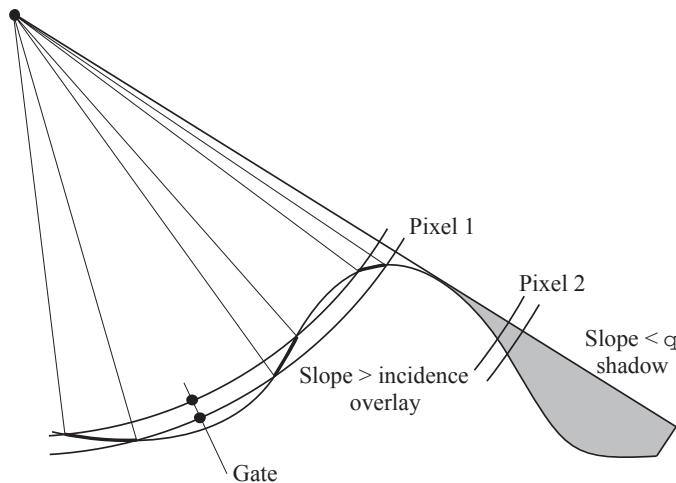


Figure 2.13. Relief effects on SAR image acquisition:
examples of zone overlapping on the same gate and radar shadow

Conversely, if the slope is negative (opposed to the radar), the gates decrease in size and the image dilates. If the slope is too steep ($\alpha > \theta - 90^\circ$), the area is not illuminated and no energy is scattered back to the radar: this is called radar shadow. Such hidden areas are frequent during acquisitions by airborne radars, as incidences in this case are typically higher than those used in spaceborne radars.

This page intentionally left blank

Chapter 3

Existing Satellite SAR Systems

What progress has been made so far in processing SAR images is largely due to the development of civilian satellite radars ERS and RADARSAT that have helped spread images among the various remote sensing communities. However, numerous other systems have been launched in the past 30 years. While most of them have remained classified as defense-related, several others have made it possible to test the broad diversity of radar imagery and its applications by proposing various resolutions, frequencies and polarizations.

Table 3.1 is a basic characteristic recap of the five main families of past and current spaceborne SAR acquiring commercially available images.

We will examine these systems at length in this and other chapters, but first we will cover fundamentals that will make it easier to understand the orbitography of such satellites.

3.1. Elements of orbitography

The main characteristic of regular Earth coverage by imaging satellites depends on their particular type of orbit (orbital plane inclination, phasing, cycle, etc.). This section provides complementary information on some orbitography notions that can improve our understanding of several essential parameters of SAR systems and our approach to some satellite location issues. Several acquisition elements playing a

major role, for example, in the implementation of relief reconstruction techniques are also specified.

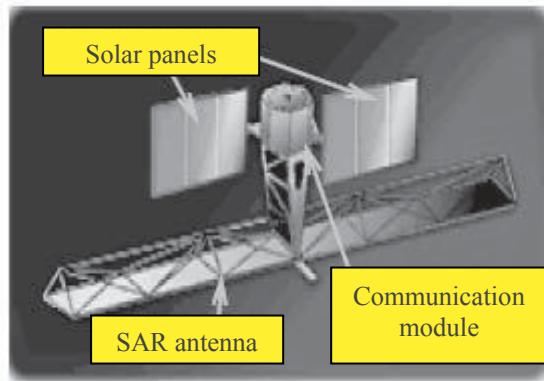


Figure 3.1. Example of a spaceborne SAR: the RADARSAT satellite © CSA

3.1.1. Remote sensing satellite orbits

Among the numerous parameters characterizing orbits we will only discuss those that are relevant to Earth observation satellites. Such satellites must first of all have near-polar orbits in order to ensure their coverage is as fully global as can be. As for altitude, most of them operate at around 800 km. They are often sun-synchronous, although this requirement is mainly useful for optical remote sensing satellites.

3.1.1.1. Inclination

Orbital inclination is the angle between the equatorial plane and the orbital plane of the satellite in ascending direction (i.e., when it is headed for the North Pole; see Figure 3.2). For sun-synchronous satellites, this angle has to be higher than 90°, as will be shown below. It is actually 98.5° in ERS and 98.6° in RADARSAT.

3.1.1.2. Period

The velocity of a satellite in a circular orbit depends only on its altitude H :

$$V_{sat} = R_E \sqrt{\frac{g}{R_E + H}}. \quad [3.1]$$

Its orbital period, which is the time it takes to travel one full orbit, depends on its altitude H and the Earth's radius R_E as follows:

$$T_{sat} = 2\pi \sqrt{\frac{(R_E + H)^3}{gR_E^2}}$$

Regarding remote sensing satellites, the altitude frequently ranges from 500 to 800 km: they have orbital periods of 100 min, approximately 14 orbits per day.

	ERS-1&2	RADARSAT	JERS-1	SEASAT	ALMAZ
Band	C	C	L	L	S
Polarization	VV	HH	HH	HH	HH
Frequency	5.3 GHz	5.3 GHz	1.275 GHz	1.275 GHz	3.125 GHz
Wavelength	5.66 cm	5.66 cm	23.53 cm	23.5 cm	9.6 cm
Antenna	10 m × 1 m	15 m × 1.5 m	12 m × 2.5 m	10.74 m × 2.16 m	15 m × 1.5 m (× 2)
B	15.55 MHz	11.6; 17.3; 30.0	15 MHz	19 MHz	
τ	37.12 μs	20 μs	35 μs	33.4 μs	50 ns
Sampling	18.96 MHz	14–32 MHz	17.076 MHz	19 MHz	38 MHz
PRF	1,640–1,720 Hz	1,200–1,400 Hz	1,500–1,600 Hz	1,463–1,640 Hz	1,500 Hz
Look angle	23°	20° to 49°	35°	19° to 26°	32° to 65°
Swath	100 km	100 to 400 km	75 km	100 km	40 to 280 km
No. of looks	3	4	3	4	> 4
Resolution	20 m	9–100 m	18 m	25 m	15–30 m
Altitude	782 km	797.9 km	568 km	800 km	270–380 km
Period	100 min	100.7 min	96 min	101 min	92 min
Orbits per day	14+11/35	14+7/24	15	14	
Inclination	98.542°	98.594°	97.6°	108°	72.7°
Sun-synchronous	yes	yes	yes	no	no
Ascending node	22:30	18:00	22:45	***	***
Descending node	10:30	06:00	10:45	***	***
Cycle (no. of orbits)	35 days	24 days	44 days	17.3 days	1–4 days
Weight (tons) including propellants	501	343	660	2.29	18.5
	2.4	2.75	1.4		1,300 kg
57 kg					
Launch date	07/17/1991 04/20/1995	11/04/1995	02/11/1992	06/28/1978	03/31/1991
End of transmissions	03/10/2000		10/11/1998	10/10/1978	10/21/1992

Table 3.1. Main characteristics of commercial satellite SAR systems

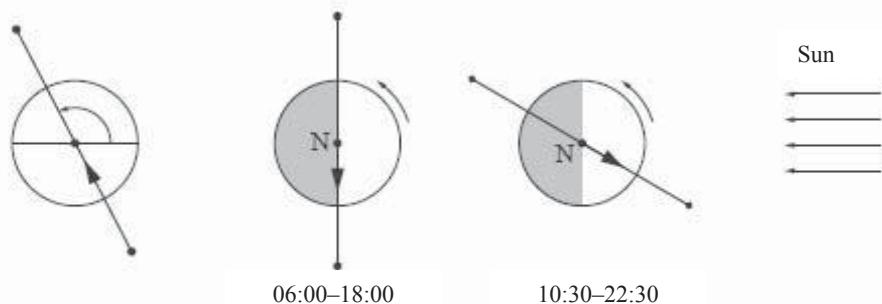


Figure 3.2. Left: conventional orbit representation with respect to equatorial plane.
Right: orbit examples: RADARSAT (06:00–18:00) and SPOT (10:30–22:30)

3.1.1.3. Eccentricity

The ellipticity of an orbit is characterized by its eccentricity: $e = \sqrt{\frac{a^2 - b^2}{a^2}}$,

where a and b represent semi-major and semi-minor axes of the orbit. Remote sensing satellites often have quite circular orbits (i.e., their eccentricity is very low, e.g., for ERS, $e = 1.165 \cdot 10^{-3}$).

3.1.1.4. Sun-synchronism

A satellite is said to be sun-synchronous (or heliosynchronous) when it is synchronized with the Sun, i.e., when there is a constant angle between its orbital plane and the Earth-Sun axis. In these conditions, the satellite benefits from nearly identical illumination for every scene that it observes (except for latitude and seasonal variations). While being sun-synchronous has no relevance for radar satellites, they are often so. As a result, they can use the same platforms (solar panel orientation) as their optical counterparts and have a simpler operating pattern. ERS 1&2 satellites for example, are referred to as “10:30–22:30” like the SPOT satellites whose platform they share (see Figure 3.2). This means they cross the equator (descending node) at 10:30 in the morning (local sun time) and again (ascending node) at 22:30. They go down on the sunlit side of the Earth and go back up on the shadowed side (during the night). RADARSAT is a “06:00–18:00” satellite, meaning that it does not enjoy good ground illumination by the Sun, which it does not really need. Still, its solar panels are constantly illuminated, which allows it to perform acquisitions on both its ascending and descending orbits without using too much battery power.

Sun-synchronism is made possible by an anomaly of the Earth potential. If the Earth were a perfectly spherical homogenous body, orbital planes would be time-invariant (with respect to a reference point in the sky). The Earth's flattening at the poles is, among other things, responsible for a rotation of this plane, known as precession. By choosing orbital parameters carefully, we can make this plane spin at a rate of one turn per year to make up for the Earth's rotation around the Sun. For a circular orbit, this implies the following relation between altitude H and orbital inclination i (which has to be wider than 90° for the cosine to be negative):

$$-9.97 \frac{R_T}{R_T + H} \cos i = \frac{360}{365.24} = 0.985$$

3.1.1.5. Cycle

For the sake of comparison, it is desirable that successive images of a same region are acquired at the same look angle. We therefore want the satellite to pass over the same point on the Earth's surface once it has completed an orbital cycle involving an equal integer number of days (same Earth's position) and equal integer number of orbits (same satellite position). Such orbits are described as phased or geo-synchronized. ERS and RADARSAT cycles are 35 days and 24 days long, respectively. Phased orbits often have a sub-cycle, which is the time it takes the satellite to pass again (though not exactly with the same look angle) near a given point. The RADARSAT platform, for example, travels $(14 + 7/24)$ orbits per day. In seven days, it makes $(7 \times 14 + 7/24)$ revolutions, a nearly whole number, and is thus very close to the orbit it followed on the first day. In 24 days, it completes an integer number of orbits and begins another orbital cycle which is entirely identical to the previous one.

3.1.1.6. Phasing

Sun-synchronous orbits are often phased (existence of a cycle), which leads to some confusion between these two attributes. Still, not all sun-synchronous orbits are phased, nor are all phased orbits sun-synchronous. From a mathematical perspective, phasing implies H , but not i .

3.1.1.7. Orbital drift and correction

At an altitude of 800 km, the atmosphere is rarefied, yet friction continues to exist, albeit rather low. It leads to loss of altitude and, according to equation [3.1], acceleration. The result is an eastward drift: the Earth always spins at the same speed, so the satellite arrives in the East a little earlier than planned. The ERS, for example, loses around 1 m daily.

In order to keep orbital properties unchanged, the satellite needs to be repositioned to its nominal altitude at regular intervals. Adjustment operations are ground-controlled and involve the use of thrusters mounted on the platform and on-board propellant.

3.2. Polar orbiting SAR satellites

SAR sensors carried by satellites in polar orbits generally look to the right of the satellite. Given the orbital inclination, these side-looking sensors can image the North Pole (actually an area within a few square kilometers of it), but not the South Pole (unless the satellite is overturned, as will be shown in section 3.2.4).

3.2.1. *SEASAT*

SEASAT (SEA SATellite), the first civilian SAR system ever put into orbit, was launched in 1978 by NASA. Unfortunately, it went out of operation due to a failure of the electric system some 100 days later. Its main characteristics are given in Table 3.1.

3.2.2. *ERS*

ESA's ERS ("European Remote Sensing" satellite) 1 and 2 were placed into orbit in 1991 and 1995, respectively, and were equipped with a variety of instruments, including a C-band SAR. The satellite ERS-1 went out of service over a platform failure on March 10, 2000. Their main characteristics are recapped in Table 3.1. Some peculiarities of these twin systems deserve additional specifications.

Tandem phase

For the period between April 20, 1995 and March 10, 2000, the European Space Agency had two SARs in orbit: ERS-2 was a clone of ERS-1 and had the same orbit with a 30-minute offset. This made it possible to have the same spot on the ground flown over by ERS-1 and then revisited 24 hours later by ERS-2.



Figure 3.3. SAR spaceborne satellites: ERS-1 (left), JERS 1 (center) and Envisat (right)

Acquisition limitations

When on the sunlit side, acquisition times for ERS 1 & 2 are limited to 12 minutes per orbit (the acquisition of a scene requires about 15 seconds). It shrinks to just 4 minutes during night pass.

ERS-1 phases

Since the ERS sensor does not have an electronic focusing system that would allow a variation of its look angle, ESA defined specific phases within the ERS-1 mission in order to test different experiment configurations. These tests referred to, among other things, reorienting the system to a look angle close to 35° (roll-tilt mode) and adjusting the orbit to obtain a 3-day cycle for certain given areas of the globe.

Phases		Start	End	Cycle (days)	Altitude (km)	Inclin. (°)	Orbits
A	<i>Commissioning</i>	07/25/91	12/10/91	3	776	98.516°	43
B	<i>Ice</i>	12/28/91	04/01/92	3	776	98.516°	43
R	<i>Roll Tilt</i>	04/02/92	04/14/92	35	782	98.542°	501
C	<i>Multidisciplinary</i>	04/14/92	12/23/93	35	782	98.542°	501
D	<i>Ice</i>	12/23/93	04/10/94	3	776	98.516°	43
E	<i>Geodetic</i>	04/10/94	09/28/94	168	770	98.491°	2411
F	<i>Geodetic</i>	09/28/94	03/21/95	168	770	98.491°	2411
G	<i>Multidisciplinary</i>	03/21/95	03/10/00	35	782	98.542°	501

Table 3.2. Activity phases of the ERS-1 sensor (the columns Inclin. and Orbit indicate respectively inclination of the orbit plane and the number of orbits per cycle)

Note also that in practice the PRF is almost always 1679.9 Hz.

3.2.3. JERS

Like ERS, the Japanese Earth Resource Satellite (JERS) carries various instruments, including an L-band SAR (its characteristics are illustrated in Table 3.1). Managed by NASDA, the National Space Development Agency of Japan, it was placed in a polar orbit (568 km altitude and 97.7° inclination) in June 1992. Its transmissions ended following an electric failure in October 1998.

3.2.4. RADARSAT

The RADARSAT (RADAR SATellite) system of the Canadian Space Agency (CSA) is based on an antenna which consists of 8 identical panels (in azimuth direction), each divided into 32 sub-antennae that help adjust look angle via phase control. This electronic control of the look angle is an original RADARSAT feature. Note, however, that transmitting and receiving differ in phase shift values. As a result, phase shift values need to be adjusted between the two modes (the electronic gear can operate the change in less than 20 μ s). Seven angle values are available in standard mode (mode 1). The physical antenna looks at 30° (approximately beam 3 in mode 1). On the other hand, this electronic pointing system enables it to work in several specific modes (e.g., Wide, Scan SAR, etc.).

Table 3.1 provides the main characteristics of this sensor.

Note that a larger azimuth than ERS-1 (15 m compared with 10 m) was chosen at some cost to azimuth resolution (rated as 9 m in single-look). In practical terms, this implies a lower PRF (pulse repetition frequency; see section 2.1.2), 1,200 to 1,400 Hz, instead of 1,640 to 1,720 Hz, leading to a longer analysis of the received trains. The swath therefore can be wider (which is possible due to electronic focusing), up to a maximum theoretical value of 125 km (at 1,200 Hz). However, we should note that the duration of the desired signal will become shorter as we subtract transmission time (20 μ s) and also the time it takes to switch the antenna from transmit to receive mode (20 μ s).

Antarctic RADARSAT mode

RADARSAT was designed to allow the satellite to rotate in orbit to 180° and observe Antarctica. This imaging mode was in operation from September 12, 1997, to November 4, 1997 and helped perform the first ever mapping of the continent.

Acquisition modes

RADARSAT is equipped with an electronic antenna pointing system which accommodates a considerable number of acquisition modes. An illustration of these different modes is given in Figure 3.4.

The products that can be made available (by choosing some specific values of the angle of incidence and other parameters) are listed in Table 3.3.

The bandwidth varies; it is 11.58 MHz (S3 to S7), 17.28 MHz (S1 and S2), and 30 MHz in Fine Resolution Mode.

3.3. Satellites in non-polar orbit

3.3.1. ALMAZ

Launched in 1991, the SAR sensor ALMAZ-1 was fitted in an available structure of the military space station ALMAZ (Russian for “diamond”) from which the Salyut stations were derived in the 1970s. This explains its mass and its capacity of carrying significant amounts of residual propellants that allowed it to change orbits. An ALMAZ prototype was loaded on Cosmos 1870 in 1987 and supplied radar images of 30 m resolution for the next two years.

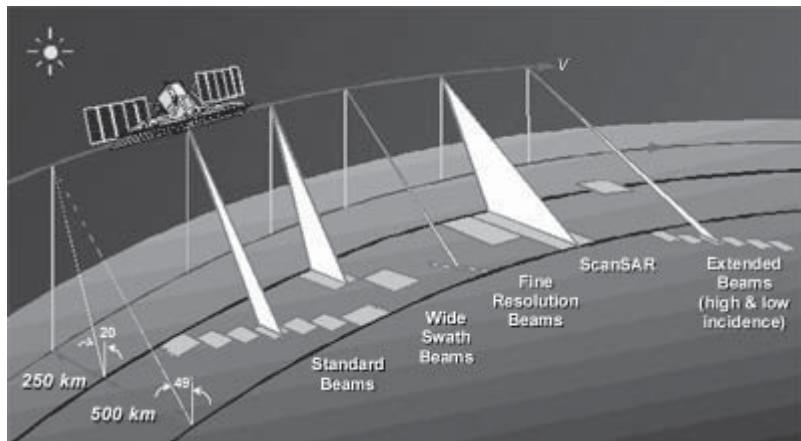


Figure 3.4. RADARSAT acquisition modes © CSA (according to [ANO 93])

Mode	Elevation × azimuth (m × m)	Number of looks	Swath width (km)	Look angle (degrees)
Standard	25 × 28	4	100	S1: 22 S2: 27 S3: 33 S4: 36 S5: 39 S6: 44 S7: 47
Wide - 1	48–30 × 28	4	165	20–31
Wide - 2	32–25 × 28	4	150	31–39
Fine resolution	11–9 × 9	1	45	F1: 38 F2: 40 F3: 42 F4: 44 F5: 46
ScanSAR narrow	50 × 50	2–4	305	20–40
ScanSAR wide	100 × 100	4–8	510	20–49
Extended (H)	22–19 × 28	4	75	50–60
Extended (L)	63–28 × 28	4	170	10–23

Table 3.3. RADARSAT system modes

ALMAZ is the only radar that has operated at the S-band so far. Images were either recorded on board or transmitted in real time via a relay satellite. Range pulse compression was not available. Transmitted pulses had about 50 ns duration with 290 kW peak power. Sampling was carried out onboard in real rather than complex mode, which resulted in a PRF of 3,000 Hz for an azimuth bandwidth of 1,500 Hz.

3.3.2. Space shuttle use

The US shuttle has carried out some interesting SAR imaging experiments, including in particular the multi-sensor multi-polarization SIR-C/X-SAR experiment and very recently the SRTM mission that used two antennae to collect interferometric data over 80% of the Earth's landmass and develop a digital elevation model of quality. The limitations of such programs involve reduced time

coverage, as shuttle missions are inevitably shorter than satellite ones, and an exploration range is confined to northern and southern latitudes less than about 60°.

3.3.2.1. Columbia: the SIR-A program (1982)

The Shuttle Imaging Radar A (SIR-A) was flown aboard the space shuttle Columbia, launched on November 12, 1982. The overall characteristics of the system are given in Table 3.4.

	SIR-A	SIR-B
Altitude	259 km	360; 257; 224 km
Inclination	57°	57°
Band	L	L
Polarization	HH	HH
Frequency	1.275 GHz	1.275 GHz
Wavelength	23.5 cm	23.5 cm
Bandwidth	6 MHz	6 MHz
Look angle	50° +/- 3°	15° to 65°
Resolution	40 m	20 to 30 m (4 looks)
Swath width	50 km	20 to 40 km
Antenna	9.4 m x 2.16 m	10.7 m x 2.16 m
τ	30.4 μ s	30.4 μ s
PRF	1,464 to 1,824 Hz	
Look equivalent		4

Table 3.4. Main characteristics of SIR-A and SIR-B sensors

Despite its short 2.5-day length, the mission revealed the high potential of these sensors, especially in the exploration of desert areas, where radar waves (L-band: $\lambda = 23.5$ cm) penetrated below the surface and helped uncover ancient dried up rivers.

3.3.2.2. Challenger: the SIR-B program (1984)

The space shuttle Challenger, launched on October 5, 1984 (flight mission STS-41-G), carried the Shuttle Imaging Radar B (SIR-B), designed to take into account earlier missions SIR-A and SEASAT. Its major characteristics are provided in Table 3.4. The system used a mechanical tilt to change the look angle between 15° and 65°. The mission, which lasted some 8 days in all, unfolded over three phases at three different altitudes, the last of which provided a repeat cycle of about 1 day.

3.3.2.3. Endeavour: the SIR-C program (1994)

With a multi-frequency multi-polarization SAR system in its cargo bay, space shuttle Endeavour completed two 10-day flights (missions SRL-1 on STS-59 in April 1994 and SRL-2 on STS-68 in October 1994). In service on both missions were the US-made Spaceborne Imaging Radar-C (SIR-C) operating at C- and L-band and the X-SAR radar built by a German/Italian team (DLR and ASI), working at X-band. Both flights took place at 225 km altitude and 57° orbital inclination.

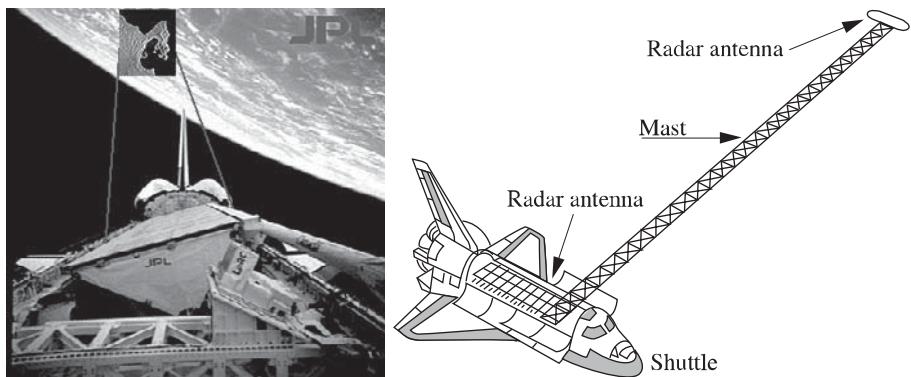


Figure 3.5. Left: SIR-C satellite; right: the shuttle Endeavour and deployed mast of the SRTM mission

Collected data were stored on board on 200 digital cassettes with a storage capacity totaling 93 Tbits. The radar system weighing 11 tons occupied three-quarters of the shuttle cargo bay.

	SIR-C	SIR-C	X-SAR
Band	L	C	X
Frequency	1.250 GHz	5.3 GHz	9.6 GHz
Wavelength	23.5 cm	5.8 cm	3.1 cm
Dimensions	12 m × 2.9 m	12 m × 0.75 m	12m × 0.4 m
Polarization	HH HV VH VV electronic	HH HV VH VV electronic	VV mechanical
τ	33.8; 16.9; 8.5 μ s	33.8; 16.9; 8.5 μ s	40 μ s
PRF	1,395 to 1,736 Hz	1,395 to 1,736 Hz	1,395 to 1,736 Hz
Look angle	20° to 55°	20° to 55°	20° to 55°

Table 3.5. The family of SIR systems

3.3.2.4. Endeavour: the SRTM program (2000)

The Shuttle Radar Topography Mission (SRTM) was implemented aboard the space shuttle Endeavour (mission STS-99) at around 233 km orbital altitude and 57° inclination during the period of February 11–22, 2000. Aside from a system similar to that of the SIR-C mission (minus the L-band sensors that had to be left behind for mass reasons), the shuttle flew a receiver-only antenna mounted at the end of a 60 m mast that was deployed in orbit at an angle of 45° from the vertical (see Figure 3.5). The system allowed interferometric data to be collected simultaneously by the external and internal antennae in response to signals transmitted by the internal radar. The SRTM mission covered land between -56° and +60° latitudes, or about 80% of the Earth's landmass [WER 00].

Furthermore, the C-band used a Scansar mode, its expanded swath width allowing global coverage that the X-band was unable to ensure.

The mission was sponsored by the National Imaging and Mapping Agency (NIMA), formerly known as the Defense Mapping Agency (DMA), DLR (Deutsches Zentrum für Luft und Raumfahrt) and, for X-band related matters, ASI (Agenzia Spaziale Italiana).

3.4. Other systems

From among the other imaging SARs placed in orbit so far, the following are worth mentioning:

- Cosmos 1870 (actually ALMAZ-K), launched on July 25, 1987 and equipped with an S-band SAR with 9.6 cm wavelength and 30 m resolution;
- Resurs-O1 N2 (1988) and Resurs-O1 N3 (1994), which orbited at altitudes of 650 to 670 km with an inclination of 98°. These were L-band (1.286 GHz bandwidth) SARs having a resolution of 50 m and transmitting at a look angle of 35°;
- TRAVERS SAR carried by the Priroda module of MIR station and operating at L-band (23 cm) and S-band (9.3 cm) with 150 m resolution at a look angle of 35°. With 50 km swath width, the system produced its first and apparently last images on September 9, 1996. Indeed, it seems the mission was affected by the troubled career of MIR itself, although a scientific meeting of Priroda Project on June 16-17, 1998, recommended that the broken sensor should be fixed;

– Lacrosse series of US military SARs, (its resolution is believed to be metric or even submetric). Three such systems are currently in orbit:

- Lacrosse 1 was carried by a shuttle (flight STS-27) launched December 2, 1988, and operated at 650-700 km altitude and 57° orbital inclination. Its mission ended in March 1997,

- Lacrosse 2 was launched on March 8, 1991, in an orbit inclined at 68°, 660 km high,

- Lacrosse 3 Indigo, launched on October 24, 1997, orbits at 670 km altitude and 57° inclination. Its mass is said to be around 14 tons.

3.5. Airborne SARs

Most airborne SARs are military systems. Still, some of them do supply data for civilian use or as part of scientific applications and some of the images they produce can even be viewed on the web.

3.5.1. *Sandia*

Sandia National Laboratories (USA) have a radar system on board a Twin Otter that can collect images in the following bands:

Ka	35 GHz
Ku	15 GHz
X	10 GHz
VHF/UHF	125–950 MHz

3.5.2. *CCRS*

In 1986, the Canada Centre for Remote Sensing (CCRS) had an airborne SAR on a Convair 580. It operates at C (5.3 GHz) and X (9.25 GHz) bands, polarimetric in C-band. The system has a resolution of 6 m x 6 m and can operate in several different modes, including nadir, narrow swath and wide swath. This system, which is still in service, has contributed to refine RADARSAT specifications. The SAREX mission carried out in South America in 1992 and GlobeSAR (Africa and Asia, 1993) helped potential RADARSAT users familiarize themselves with the data range the new sensor provided.

Laying the groundwork for the future RADARSAT-2, full polarimetric capabilities have been added to the system.

3.5.3. CRL/NASDA

Japan's Communications Research Laboratory (CRL) and National Space Development Agency (NASDA) are working together on the development of a high polarimetric resolution SAR. This X-band SAR is equipped with an active antenna and a second, passive one.

	X-Band	L-Band
Frequency	9.55 GHz	1.27 GHz
Bandwidth	100 MHz	50 MHz
Antenna	1.05 m x 0.19 m	1.55 m x 0.65 m
Look angle	10°–75° (variable)	20°–60° (fixed)
Resolution	1.5/3 m	3/6 m
Sampling	123.45/61.725 MHz	61.7/30.9 MHz
Baseline	2.3 m	

Table 3.6. Characteristics of the interferometric system under development by CRL and NASDA

3.5.4. JPL's AIRSAR

The various antennae of JPL's AIRSAR using a DC8 as a platform have multiple capabilities, particularly in the polarimetric (POLsar) mode and interferometric modes (TOPSAR: cross-track interferometric SAR and along-track interferometry SAR, or ATI). Acquisitions at the different bands are performed simultaneously. Three bandwidth values are available (20 MHz; 40 MHz and 80 MHz). This sensor that has been in operation since 1988 has served as a test bed for the SIR-C missions.

Mode	Band	λ	Antenna (in m)	Polarization
POLsar	L	25 cm (1.25 GHz)	1.613 x 0.45	4 modes, used by ATI
	C	5.7 cm (5.29 GHz)	1.359 x 0.165	4 modes, used by ATI
	P	68 cm (0.438 MHz)	1.870 x 0.912	4 modes
TOPSAR	L	25 cm	1.829 x 0.203	
	C		1.676 x 0.165	
ATI	C	5.7 cm	1.359 x 0.165	4 modes, baseline = 1.93 m
	L		1.613 x 0.45	4 modes, baseline = 19.8 m

Table 3.7. Characteristics of AIRSAR system modes

3.5.5. PHARUS

The PHARUS (PHased ARray Universal SAR) project was developed in the Netherlands in the 1990s. It is a C-band (5.6 cm), polarimetric radar designed and built by TNO, the Dutch National Aerospace Laboratory (NLR) and the Technological University of Delft. Fitted in a Cessna, this phased-array-type sensor with a resolution of 3 m served as image simulator for the Envisat ASAR.

3.5.6. DLR's E-SAR

DLR has a multi-frequency SAR with metric resolution on board a Dornier 228. Using an additional antenna, it is capable of carrying out interferometric acquisitions at X-band.

Band	X	C	S	L	P
Band (in GHz)	9.6	5.3	3.3	1.3	0.450
Polarization	H + V	H + V	H + V	H + V	H + V
Bandwidth (MHz)	120	120	120	100	25 (60)

Table 3.8. The E-SAR system

3.5.7. Some remarkable military systems

Given the defense missions of military airborne SARs, very little information usually leaks out to the public as to what systems of this kind may be operating around the world. Note, though, that in a few cases, data is publicly available and reveals the current directions that can be pursued in terms of SAR imaging:

- RAMSES (the French acronym for multi-band airborne radar for target signature study) station has been used by the French Aerospace Research Agency, ONERA, for collecting radar images across a broad range of frequencies in different polarizations. Installed aboard a Transall C160 plane, it consists of several radar components covering different bands with submetric resolutions [LEC 98]. Several types of wave forms can be used;

Band	L	S	C	X	Ku	Ka	W
Band (in GHz)	1.6	3.2	5.3	9.6	14.3	35	95
Bandwidth (MHz)	200	300	300	600	300	600	500
Elevation aperture	23°	30°	33°	16°	14°	5°, 20°	3°, 5°, 10°, 20°
Azimuth aperture	16°	10°	8°	16°	14°	5°, 20°	3°, 5°, 10°, 20°

Table 3.9. RAMSES radar main characteristics (after [LEC 98])

– CARABAS (Coherent All RAdio BAnd Sensing) has been installed by Sweden's National Defense Research Establishment (NDRE) on board a Sabreliner. The system operates at very low frequencies (20–90 MHz; 2.5 MHz bandwidth) [HEL 96a] yielding high 3 m x 3 m resolutions. Note that such a system, if used on board a satellite platform, would have had to cope with ionosphere propagation phenomena.

CARABAS and RAMSES (operating at X- and L-bands only) jointly conducted the RAMCAR98 experiment over two forest test sites in France (Landes and Lozère) to analyze the backscattering characteristics of forests and estimate their biomass.

This page intentionally left blank

Chapter 4

Synthetic Aperture Radar Images

We have explained in earlier chapters how a SAR system works and how data directly related to local ground reflectivity can be obtained by processing the signals received by a monostatic antenna. This analysis helped us assess the resolution limits of such systems: whether or not we are able to distinguish between returns from two objects lying close to each other on the ground depends on several well determined physical parameters (antenna dimensions, platform speed, and type of the transmitted wave in terms of frequency and pulse duration).

This chapter is devoted to the images that will be used in this book and on which several types of processing will be assessed. In terms of their degree of pre-processing, images fall into several categories:

- raw data that are in fact two-dimensional data prior to their SAR processing (level 0);
- single-look complex data resulting from raw data 2-D processing: pulse compression and synthetic aperture algorithm (level 1); and
- amplitude image data that is generally multi-look (levels 2 and 3).

We will examine the problem linked to the calibration of such data and the possibility of obtaining data at regular intervals considering the geographic location of the studied region and the operating requirements imposed by satellite tracking.

Polarimetric data will be addressed separately in section 4.1.5.

Chaper written by Frédéric ADRAGNA, Sylvie LE HÉGARAT-MASCLE and Jean-Marie NICOLAS.

4.1. Image data

Space agencies and satellite image providers usually supply data in the traditional CEOS format. The Committee on Earth Observation Satellites (CEOS) was created in 1984 and groups together the world's main space agencies, such as CNES, NASA, ESA, NASDA, etc. Among other things, the committee is responsible for developing an international standard for the distribution of remote sensing data products. Space agencies issue radar data in this particular format. The CEOS format consists of several files, which generally are of different sizes; it includes "image" data as well as auxiliary data, depending on the required product type and specifications.

4.1.1. Raw data

We can hardly speak of images as far as raw data are concerned. The information they contain, although supplied in two-dimensional tables physically related to the sensor (antenna direction, satellite track), is not apparent at first glance, since it corresponds to a sensor whose range and azimuth resolutions are in the order of several kilometers, i.e. hundreds of rows and columns.

Space agencies nevertheless offer this type of data making it possible to perform a specific SAR synthesis. To this end, raw data are accompanied by additional related information (satellite telemetry, transmitted signal, etc.).

Raw data are digitized on a low number of bits. ERS data are coded on 5 bits for the real part and another 5 bits for the imaginary one. Note that the Magellan probe uses as few as 2 bits for the real part and 2 for the imaginary one, which actually leaves only one bit of information aside from the signal. Since raw data must transit as fast and effectively as possible between the sensor and a ground receiver station, a lot of studies have focused on developing techniques for the compression (or coding) of this type of data with or without losses. Some of these compression schemes have been implemented (for example, for RADARSAT).

4.1.2. Complex single-look data

Once raw data have been processed in terms of range (pulse compression) and azimuth (SAR synthesis), we obtain a two-dimensional table of complex data (antenna direction, satellite track) physically related to the sensor, but which can now be described as images: an amplitude visualization is indeed representative of the imaged area. This type of data is also referred to as Single-Look Complex (SLC).

At this stage, the dimension of a pixel depends on just two parameters:

- the azimuth PRF (pulse repetition frequency). Knowing the speed of the satellite V_S , the size of a pixel is given by the following formula (see section 2.1):

$$\Delta_a = L_{PRF} = V_S T_{PRF}$$

- the range sampling frequency. Knowing the look angle θ of the wave, we obtain on flat ground:

$$\Delta_{d, \text{flat ground}} = \frac{c}{2F_e \sin \theta}$$

ERS SLC pixel size	
Range	$\Delta_d = \frac{c}{2F_e \sin(\theta)} = 26 \text{ to } 18 \text{ m}$
Azimuth	$\Delta_a = V_S T_{PRF} \sim 4 \text{ m}$

Note that the look angle varies (by 25% for ERS-1) between the near range and far range.

The major advantage of this type of data is its wealth: aside from amplitude which can be visualized and processed like an image, phase will be the foundation of interferometric techniques, as it will be seen in Chapter 12. However, images of this kind are very speckled, which, in addition to the pixel geometry (notably, its rectangular shape is very elongated in range), makes visual interpretation tricky.

In the case of satellite data, SLC images are provided in 16-bit data (signed short), in which every pixel has two characteristic values (real part, followed by imaginary part)¹. For example, an ERS scene covering an area of 100 km x 100 km contains approximately 26,000 lines of 5,000 pixels each, resulting in a volume of 520 Mbytes².

¹ Other formats are available for airborne radars, e.g., formats in which the amplitude image and then the phase image are provided.

² Such a volume can easily be archived on a single CD-ROM; for tape archiving, the scene has to be divided into quadrants.

SLC data are oriented along the sensor, i.e. along the satellite. Images are therefore parallel to the satellite track. Angle $\tilde{\beta}$ between the latter and geographic North depends on orbit inclination as well as on the latitude of the observed point and on the Earth's rotation.

As far as ERS data are concerned, it is important to note that:

- ESA has opted for a convention which, though logical in terms of data, makes the reading of an image more difficult. It consists in considering the data layout in relation to the sensor acquisition. The lines are chronologically acquired and data on each line are also chronologically stored. The first pixel therefore corresponds to the near range, and the last one to the far range. To compare an image with ground truth, while also taking into account that the sensor is “right-looking”, several adjustments have to be made as follows:

- read the lines from right to left for the descending phase;
- read the table from bottom to top for the ascending phase;
- data are processed by a Processing and Archiving Facility (PAF). Since every PAF has its own signal processing algorithms, the data we obtain depend on how they have been processed. In multi-data analysis, it is therefore strongly advised that only data from the same PAF are used.

4.1.3. Multi-look data

In order to make their products easily usable, space agencies provide processed data (generally amplitude data), called PRrecision Images (PRI) easier to analyze and such that square pixels are based on ground geometry. To this end, two major steps are required, namely:

- azimuth filtering so that a reduction of speckle can be noted (some techniques are listed in section 6.2.2.1). During this step, sub-sampling is performed;
- range over-sampling taking into account look angle variation along the swath.

The final pixel size is traditionally a multiple or a sub-multiple of 25 m x 25 m. For ERS, the PRI pixel size is 12.5 m x 12.5 m (compared with the SLC pixel which is 20 m x 4 m).

ERS PRI pixel size	
Range	$\Delta_r = 12.5 \text{ m}$
Azimuth	$\Delta_a = 12.5 \text{ m}$

Note that data are over-sampled in range and sub-sampled in azimuth in terms of nominal resolution.

These data are no longer complex values, but rather amplitude values. Given the SAR imagery dynamics, they are coded in 16-bit data (unsigned short). For example, an ERS scene covers an area of about 100 km x 100 km requiring 130 Mbytes.

Like SLC data, the PRIs produced by ERS depend on the archiving facility, and their geometry is related to the sensor reference system rather than a geocoded system. Furthermore, ESA corrects these data from the effects linked to antenna pattern and range spreading loss.

RADARSAT's equivalent of the PRI product is called Path Image (SGF code) or Path Image Plus (SGX code). Sampling steps depend on the selected acquisition mode and equal:

- 6.25 m (SGF) and 3.125 m (SGX) in fine resolution modes; and
- 12.5 m (SGF) and 8 m (SGX) in standard and wide modes.

4.1.4. *Derivative products*

Space agencies sometimes provide georeferenced data. Images in this case are North-South oriented. ESA for instance offers:

- geocoded (GEC) data located on the ellipsoid WGS 1984 and based on the UTM projection for latitudes within -70° and $+70^\circ$, and UPS anywhere else;
- geocoded, terrain distortion-corrected data (GTC), insofar as a quality DTM³ is available on the investigated area (ESA only has data on Western Germany, but clients can provide their own DTMs).

³ DTM = Digital Terrain Model = altitude map presented in the form of an image $h(i,j)$, every pixel of which shows the height h of a point with coordinates (i,j) .

The RADARSAT counterparts of GEC and GTC are called map image (SSG code) or precision map image (SPG code). Sampling step sizes, which depend on the operating mode, are:

- 6.25 m (SSG or SPG) in fine resolution mode; and
- 12.5 m (SSG) in standard and wide modes.

4.1.5. Polarimetric data

The AirSAR airborne radars and the SIR-C/X-SAR (on board the US shuttle) of Jet Propulsion Laboratory (JPL/NASA) were in 2001 the only imaging radars providing fully polarimetric data (at P-, L- and C-bands in the case of AirSAR; L- and C-bands for SIR-C, and X-band for X-SAR). The main data formats of these instruments are as follows.

4.1.5.1. SIR-C/X-SAR data

They may be SLC, MLC (Multi-look Complex), or MLD (Multi-look Detected) types and correspond to the following data:

Data	quad-pol	dual-pol	single-pol
SLC	Yes (10 bytes/pixel)	Yes (6 bytes/pixel)	Yes (4 bytes/pixel)
MLC	Yes (10 bytes/pixel)	Yes (5 bytes/pixel)	No
MLD	No	No	Yes (2 bytes/pixel)

These are compressed data. The following equations [CHA 94] are used to uncompress them:

- for SLC quad-pol data:

$$ysca = \sqrt{\lceil byte(2)/254 + 1.5 \rceil} \cdot 2^{byte(1)}$$

$$\text{Re}(S_{hh}) = byte(3) \times ysca / 127 \quad \text{Im}(S_{hh}) = byte(4) \times ysca / 127$$

$$\text{Re}(S_{hv}) = byte(5) \times ysca / 127 \quad \text{Im}(S_{hv}) = byte(6) \times ysca / 127$$

$$\text{Re}(S_{vh}) = byte(7) \times ysca / 127 \quad \text{Im}(S_{vh}) = byte(8) \times ysca / 127$$

$$\text{Re}(S_{vv}) = byte(9) \times ysca / 127 \quad \text{Im}(S_{vv}) = byte(10) \times ysca / 127$$

- for MLC quad-pol data:

$$\begin{aligned}
qscsa &= \lceil \text{byte}(2)/254 + 1.5 \rceil \cdot 2^{\text{byte}(1)} \\
|S_{hv}|^2 &= qscsa \times \left[(\text{byte}(3) + 127)/255 \right]^2 \\
|S_{vv}|^2 &= qscsa \times \left[(\text{byte}(4) + 127)/255 \right] \\
|S_{hh}|^2 &= qscsa - |S_{vv}|^2 - 2 \cdot |S_{hv}|^2 \\
\text{Re}(S_{hh} \cdot S_{hv}^*) &= 0.5 \times qscsa \times \text{sign}\{\text{byte}(5)\} \cdot \left[\text{byte}(5)/127 \right]^2 \\
\text{Im}(S_{hh} \cdot S_{hv}^*) &= 0.5 \times qscsa \times \text{sign}\{\text{byte}(6)\} \cdot \left[\text{byte}(6)/127 \right]^2 \\
\text{Re}(S_{hh} \cdot S_{vv}^*) &= qscsa \times \left[\text{byte}(7)/254 \right] \\
\text{Im}(S_{hh} \cdot S_{vv}^*) &= qscsa \times \left[\text{byte}(8)/254 \right] \\
\text{Re}(S_{hv} \cdot S_{vv}^*) &= 0.5 \times qscsa \times \text{sign}\{\text{byte}(9)\} \cdot \left[\text{byte}(9)/127 \right]^2 \\
\text{Im}(S_{hv} \cdot S_{vv}^*) &= 0.5 \times qscsa \times \text{sign}\{\text{byte}(10)\} \cdot \left[\text{byte}(10)/127 \right]^2
\end{aligned}$$

– for MLC dual-pol data: the compression formats are the same as for MLC quad-pol, except that not all bytes are present: those corresponding to unsupplied polarizations are missing;

– for MLD single-pol data:

$$P = \lceil \text{byte}(2)/254 + 1.5 \rceil \cdot 2^{\text{byte}(1)}$$

4.1.5.2. AirSAR data

AirSAR data are quad-pol data. They are essentially provided in three formats:

- complex scattering matrix data (10 bytes/pixel), which is a format identical to the SLC format of SIR-C except for one compression factor that can be found in file headers;
- Stokes matrix data format (10 bytes/pixel), which should not be mixed up with the MLC format of SIR-C;
- detected data, an uncompressed format (with pixels coded in 4 bytes real data) consisting of one file for each polarization.

From compressed data format, the following equations [VAN 91] are used to retrieve the Stokes matrix terms:

$$\begin{aligned}
 M_{11} &= \text{gen_fac} \times [\text{byte}(2)/254 + 1.5] \cdot 2^{\text{byte}(1)} \\
 M_{12} &= M_{11} \times \text{byte}(3)/127 \\
 M_{13} &= \text{sign}(\text{byte}(4)) \times M_{11} \times \text{byte}(4)/127 \\
 M_{14} &= \text{sign}(\text{byte}(5)) \times M_{11} \times \text{byte}(5)/127 \\
 M_{23} &= \text{sign}(\text{byte}(6)) \times M_{11} \times \text{byte}(6)/127 \\
 M_{24} &= \text{sign}(\text{byte}(7)) \times M_{11} \times \text{byte}(7)/127 \\
 M_{33} &= M_{11} \times \text{byte}(8)/127 \\
 M_{34} &= M_{11} \times \text{byte}(9)/127 \\
 M_{44} &= M_{11} \times \text{byte}(10)/127 \\
 M_{22} &= M_{11} - M_{33} - M_{44}
 \end{aligned}$$

4.1.5.3. Polarimetric synthesis and examples

The main difference between fully polarimetric radars and those acquiring multi-channel data (e.g., multi-frequency, multi-incidence, multi-temporal data) is that the former acquire exhaustive information (in terms of polarimetry). Indeed, once the complex backscattering **S** matrix or the Mueller **M** matrix is available, it is possible to simulate, using equation [1.20], data acquired using any transmitting or receiving polarization configuration. Polarimetric synthesis does not provide new information itself, but makes it possible to emphasize some characteristics of an imaged target, or to characterize some backscattering mechanisms. For instance, [TOU 92] uses the minimum and maximum polarization degrees to determine the type of backscattering mechanism that is in place and its degree of heterogeneity.

Figures 4.1a, 4.1b, and 4.1c exemplify the measured or simulated powers in different configurations, namely hh, hv, and vv, in which the first letter denotes transmit polarization and the second letter denotes received polarization. We can note, at least qualitatively, the complementarity between different polarizations, particularly between copolar (hh or vv) and cross-polar (hv). Figure 4.1d shows the corresponding maximum polarization degree.

4.2. Radiometric calibration

As in all measuring systems, the physical quantities to be measured are subject to distortion by the instruments (in our case, the SAR antenna in transmitting and receiving modes) and the transmission channel. Data acquired must be calibrated.

4.2.1. *Scalar SAR data calibration*

In the case of scalar data, we can establish from radar equation [2.1] the following relation between the received intensity signal s and the backscattering coefficient σ^0 (sigma nought), which represents the radar cross-section to the considered pixel:

$$s = \sigma^0 \frac{\sin \theta_{ref}}{\sin \theta} C(\theta) K_c K_{AD} K_P$$

where:

$C(\theta)$ is a quantity related to antenna pattern;

K_c is a quantity related to transmitted chirp power;

K_{AD} is a gain factor needed for digitization;

K_P is a quantity related to geometric attenuation;

$\frac{1}{\sin \theta}$ is a factor accounting for look angle (in flat ground assumption).

Providers of radar images (whether SLC or PRI) generally use precise calibration procedures allowing users to process them by considering the values as either having the physical dimension of a radar cross-section (absolute calibration) or being pixel radar reflectivity. It is thus possible to compare values from one acquisition to the next, which is important for multi-temporal analyses.

Calibration involves complex processes that require test targets and areas in order to estimate the effective pattern of the antenna as effectively as possible. As far as ERS data are concerned, an ESA document [LAU 98] underlines their complexity and the fact that preprocessing depends on PAF (see section 4.1.2) and on processing data.

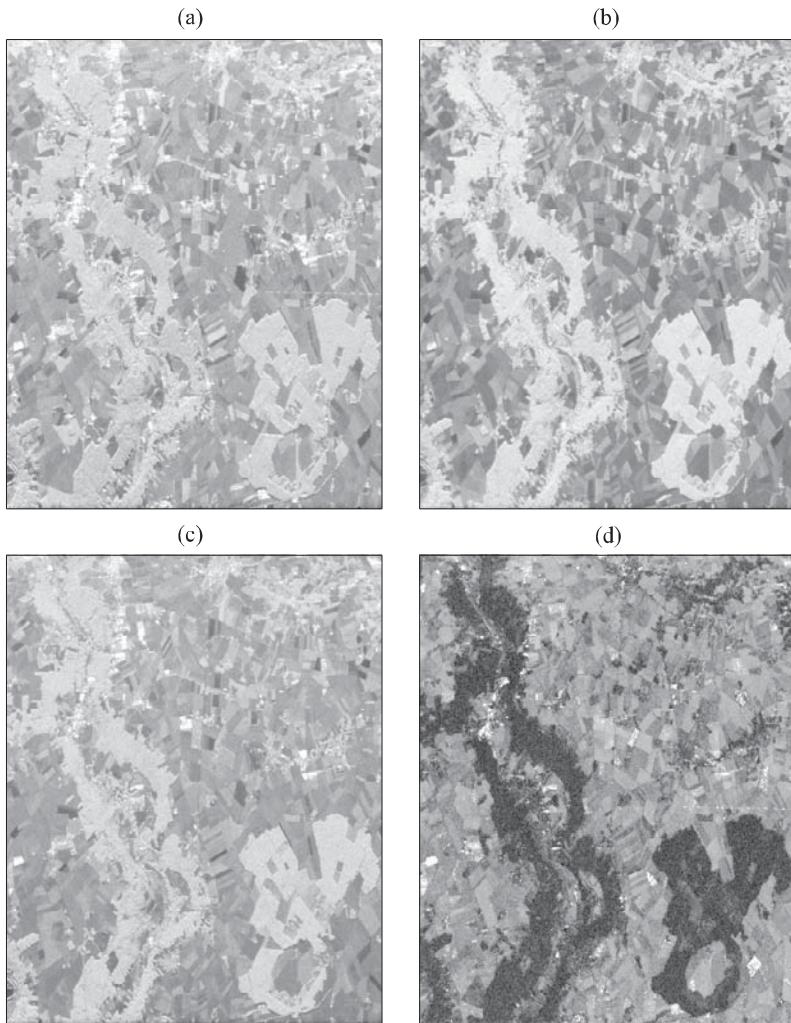


Figure 4.1. Different polarimetric configurations of the same site (Orgeval, France):
 (a) hh , (b) hv , (c) vv and (d) maximum polarization degree

4.2.2. Calibration of SAR polarimetric data

In the polarimetric data case, we consider that the received signal, which is the complex backscattering \mathbf{O} matrix, can be modeled as follows:

$$\mathbf{O} = A e^{i\varphi} \cdot \mathbf{T} \cdot \mathbf{S} \cdot \mathbf{R} + \mathbf{N} \quad [4.1]$$

where \mathbf{S} is the researched scattering matrix, \mathbf{T} and \mathbf{R} are the transmitting and receiving distortion matrices, A is an absolute amplitude factor, φ is an absolute phase and finally \mathbf{N} is the additional (thermal) noise matrix generally assumed to be negligible.

\mathbf{R} and \mathbf{T} depend on the along-track (range) distance to the considered pixel, but not on cross-track (azimuth), since SAR processing does not vary in azimuth direction. The whole calibration problem therefore boils down to estimating these matrices on every SAR image line, as well as A and φ , and to reversing equation [4.1] to obtain \mathbf{S} .

To this end, each of the \mathbf{R} and \mathbf{T} matrices is decomposed into a product of matrices, one being a diagonal matrix. Using this form allows us to distinguish between:

- the distortions due to the coupling of cross polarizations H and V in transmitting and receiving (cross-talk) that affect both the cross- and co-polar channels;
- the distortions caused by differences on the transmitting and receiving chains of channels H and V (channel imbalance) that affect the co-polar channels (gain and phase); and
- a bias on absolute gain.

Several algorithms have been proposed [VAN 90, KLE 92, FRE 92b] to remove cross-talk by using only information contained in observations and making assumptions such as the azimuthal symmetry of targets and the decorrelation of co-polar and cross-polar channels. These assumptions are generally verified as we consider a region (generally a set of image lines) containing a large enough number of natural surfaces. By contrast, channel imbalance correction and absolute calibration (in gain and phase) require a set of known response targets on the ground. These are generally manmade targets (trihedrals or dihedrals) specially deployed for calibration.

4.3. Localization precision

Auxiliary files which make it possible to locate data on the ground surface are associated with image data. The interpretation of these data, which presents many geo-referencing problems, will be discussed in Chapter 9, which is dedicated to the geometric aspect of SAR data.

This page intentionally left blank

Chapter 5

Speckle Models

5.1. Introduction to speckle

A radar imaged region, even if its physical backscattering characteristics are stationary (a “homogenous” region, as the radar people call it), appears to be made up of pixels. Their values are widely spread instead of being fairly constant, particularly in images with a small number of L looks. In single-look images, the field and backscattered power value with the highest probability is most likely to be zero! Moreover, pixel variance increases with mean radar reflectivity. As a result of such dispersion, radar images have a grainy, extremely noisy appearance, as seen in Figure 5.1. This phenomenon is due to speckle, which is also referred to as “multiplicative noise” in view of its specific properties. If all we have is just one image, speckle makes it very difficult to resolve a number of problems such as detecting objects that are small or low-contrast compared to their environment, or detecting boundaries, or discriminating surfaces. The same difficulty confronts both automated techniques and photo interpretation. Thus, to make any use of a radar image, we may have to either reduce the fluctuations of pixel values beforehand or apply processing adapted to its characteristics for any specific piece of information we desire to extract. Speckle reduction is commonly known as “filtering”, an ill-suited name, since speckle is not really noise, as will be seen further below, but rather contains information on the sensor and observed surface.

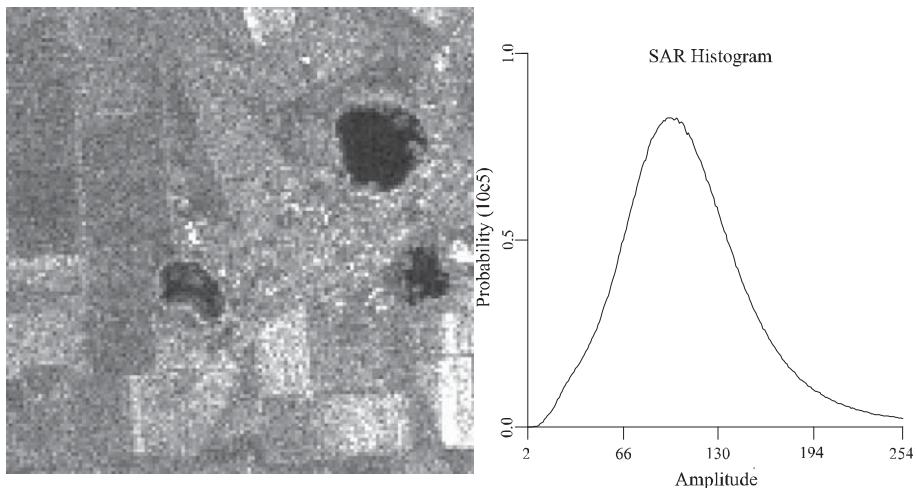


Figure 5.1. Example of a strongly magnified SAR image of an area of fields and ponds.

The speckle that affects the signal is quite clearly illustrated. On the right-hand side, a histogram of the gray levels in the area shows a continuous spreading of magnitudes from the darkest levels (the ponds) to the lightest ones (the bare fields on the right).

ERS satellite image from Ukraine (© CNES)

Under the filtering approach, historically the first to be developed, the filtering step was to be followed by the application of image processing methods, originally designed for incoherent imaging, without major changes. The main “filters” adapted to SAR images will be reviewed in Chapter 6. They have their limitations and yield less than optimal results when we try to extract specific information.

Yet, if we think of filtering as a sequence of decisions and estimates of unknown surface parameters, we see it fits perfectly well in a second approach. The concepts it provides can help develop some effective tools that can in turn be wisely used at a later, more appropriate stage during processing rather than as a pre-processing procedure. Chapters 7 and 8 will show how different processes such as detection, segmentation, classification and relief extraction can be adapted to individual characteristics and the extraction of specific parameters.

In either case, we will have to define the surface parameters used and determine the influence of speckle on our observations and consequently on their optimal assessment. Speckle modeling in this respect is therefore very important. In addition, advanced modeling makes it possible to simulate some realistic SAR images on which image processing and analysis algorithms can be validated [RAN 88]. Such modeling can be based on the physical modeling of interactions between a quasi-

monochromatic electromagnetic wave and a random environment. This particular approach however, is somewhat out of context in an image processing book such as this. For more information, see [ISH 78, BEC 63]. In addition, there are numerous books and articles devoted to physical modeling used to calculate the “mean” backscattering properties of surfaces [ULA 81, ULA 82, ULA86b, TSA 85, FUN 94, ULA 90]. Speckle is then considered as a fluctuation of observations in relation to these mean properties [ULA 86a]. Because of this, it is desirable that these fluctuations be modeled statistically.

We can thus produce adequate statistical models that will more easily fit into broader signal and image processing methods. Scattering by a set of point targets (see section 1.2.2.3) will allow us to model speckle as interference between different backscattered waves. Scattering elements are randomly spread on the resolution surface (or in the volume in the case of vegetation, snow and very dry soils) delimited by the impulse response of the sensor (or wave penetration in the case of scattering volumes). This physical phenomenon, perfectly deterministic at pixel scale, becomes random once its successive realizations on rough surfaces are observed. Statistically, it can then be described as “random walk” in the $\{xy\}$ plane. Under certain conditions vastly verified in practice, this random process results in the Gaussian or “fully developed” speckle which can be found on homogenous surfaces.

We will first describe and model this Gaussian speckle in both one- and multi-dimensional cases. This will allow us to statistically describe several variables (e.g., radar reflectivity, coherence, phase difference, polarization or channel ratio, polarimetric covariance or Stokes matrices, etc.) used in this work. Extending the model to non-homogenous surfaces, where the speckle is no longer Gaussian, will be examined in sections 5.2 and 5.3. For a low number N of scatterers, speckle is not developed and a deterministic response is obtained where $N = 1$. These elements are covered in section 5.4 since very high spatial resolution sensors operating at the lowest frequencies are necessary in this case. Finally, polarimetric aspects are discussed in section 5.5. This type of data provides the most general, complete and information-rich physical account for any given spectral band in monostatic configuration.

5.2. Complex circular Gaussian model of single-look and scalar speckle

This section will help establish the marginal distributions expression of different variables arising from complex or detected single-look data for an isolated pixel. The joint distributions of several pixels and of multi-channel multi-look data are discussed in section 5.3.

5.2.1. Random walk and Gaussian speckle conditions

5.2.1.1. Random walk in a complex plane

Consider a given polarization channel (e.g., transmit polarization h , polarization v) and denote it by (ij) . The complex amplitude z of a pixel with coordinates (x_0, y_0) is either equal to the S_{ij} component of the scattering matrix if we work with calibrated data, or proportional to the total backscattered field if our data are not calibrated. The statistical field distribution is the convolution of the complex field probability density $\varepsilon(x,y)$ with the normalized impulse radar response $h(x,y)$ (see equation [2.9]) – a dimensionless quantity. However, within a proportionality constant depending on the radar equation, z will be written $z(x_0, y_0) = \varepsilon(x,y) * h(x,y)$. This integral can be discretized assuming that z is actually the response of a finite number of quasi-point basic scatterers N_d , also called bright spots or diffractors. Then the complex field density $\varepsilon(x,y)$ is a sum of Kronecker distributions $\epsilon(x,y) = \sum_n e_n \delta(x - x_n, y - y_n)$. To simplify the coordinates, we can substitute the unique variable $u = (x,y)$ in the above relation; the measure will then be written:

$$z(u_0) = \sum_{n=1}^{N_d} e_n (u_n) h(u_0 - u_n)$$

Following weighting by $h(u)$ and other radar system parameters, z_n contribution by n scatterer has an amplitude a_n and a phase ϕ_n . If we omit the u coordinates, the complex value of the considered pixel is:

$$z = A e^{j\phi} = \sum_{n=1}^{N_d} a_n e^{j\phi_n},$$

Total field $z = A e^{j\phi}$ can be represented as a Fresnel construction in the complex plane where it results from the vectorial sum of N_d vectors (also known as phasors) representing N_d basic contributions (see Figure 5.2). In the case of speckle, if a_n and ϕ_n are random variables due to the distribution of electromagnetic properties and scatterer positions, then z is also random. In this case, the tip of vector z is found at the end of a random walk corresponding to the end-to-end sum of N_d random vectors e_n .

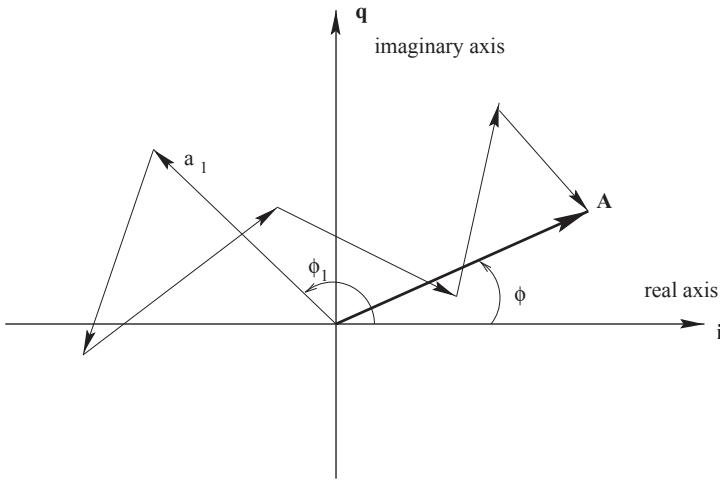


Figure 5.2. Random walk in the complex plane resulting from the addition of $N_d = 6$ scatterers

5.2.1.2. Fully developed speckle hypotheses

The complex circular Gaussian model of fully developed speckle is based on the following assumptions (see, for example, [DAI 84]):

- 1) the number of scatterers N_d per resolution cell is large enough;
- 2) the amplitude and phase of scatterer n do not depend on those of other scatterers;
- 3) the amplitude and phase of any one scatterer are independent of each other;
- 4) a_n variables are identically distributed with first and second order moments equal to $E(a)$ and $E(a^2) = \sigma$ respectively; and
- 5) phases ϕ_n are uniformly distributed between $[-\pi, \pi]$, i.e., the surface is considered rough with respect to the wavelength λ .

The first condition is met if spatial resolution is considerably higher than the wavelength λ and there is more than one object in the resolution cell. The second criterion stems from the physically simplified wave-matter interaction in which interactions between scatterers are ignored. Statistically, such simplification may be admitted if we consider clusters of electromagnetically coupled elements as forming a unique scatterer. The third criterion is based on the assumption of a randomly structured scattering medium, which is always the case with natural surfaces and even with many manmade surfaces. The fourth criterion is also the result of the

rough nature of the surface. However, even if the total response were the result of N_d identical scatterers, the interference phenomenon and statistical model thereof would still be valid provided that the first and last criteria are verified. This consideration is amply substantiated if every basic phase ϕ_n is regarded as the sum of several phase shifts: $\phi_n = \varphi_n + \arg h(u_n) + \frac{-4\pi D_n}{\lambda}$ where:

- the phase shift φ_n occurs during the physical process in which the incident wave is backscattered onto the scatterer: $\varphi_n = \arg S_{ii}$;
- the phase shift $\arg h(u_n)$ corresponds to the radar system;
- the last phase shift corresponds to the wave's two-way trip, as transmitted by the radar system and then backscattered towards the radar, which is here formulated for the case of a homogenous atmosphere and a geometric range D_n (this phase is generally associated with the optical thickness of the transmitted and reflected wave paths). This phase term is the one used in interferometry.

Remember that the Rayleigh criterion, $\Delta h > \frac{\lambda}{8\cos\theta}$ (equation [1.12]), where Δh is the standard deviation of the scatterers' height and θ is the angle of incidence, allows us to describe a surface as being rough in terms of λ and thus to satisfy criteria 5) above. In practice, the differences between the various ranges D_n in the order of $\lambda/2$ are enough to cause destructive interferences.

5.2.2. First order distributions: R radar reflectivity

5.2.2.1. Fully developed complex speckle distribution: radar reflectivity

We suppose in this section that N_d and moment $E(a^2)$ are constant. Amplitudes a_n themselves may ultimately be perfectly constant. Fluctuations in this case will only be due to the random distribution of scatterers and, to a lesser extent, to a random draw on scatterers inside the resolution cell.

We will first consider the case of a single pixel and simply imagine the scatterers undergoing slight random motions, as we have seen, of $\pm\lambda/4$, so that the observed resolution cell and its electromagnetic properties are not disrupted. The spatial extension of this hypothesis would correspond to an ideal, perfectly homogenous surface where the cells would contain exactly the same scatterers but in a different layout. In a more realistic case, where the number of scatterers would still be constant and belong to the same statistic population $\{a\}$, but would change from one pixel to the next that we will call homogenous. Finally, if the number of scatterers N_d and the statistic parameters such as $E(a^2)$ both vary, the result is a so-called heterogenous surface that will be discussed in section 5.4. What we are only

interested in this section is the marginal distribution of z and the derived variables. To simplify our notations, we will also consider an impulse response of the gate function type in amplitude $|h(u)| = I_{\delta u}(u)$ over a resolution width δu . The effect of $h(u)$ is especially involved in the spatial characterization of speckle (2nd order statistics), as will be seen in section 5.3 below. The statistical distribution of z is found by taking into consideration its real part denoted i (for “in phase”) and its imaginary part denoted q (for “quadrature”) in complex plane z ($z = i + jq$). They are obtained from:

$$\left\{ \begin{array}{l} i = \sum_{n=1}^{N_d} a_n \cos \phi_n, \\ q = \sum_{n=1}^{N_d} a_n \sin \phi_n. \end{array} \right.$$

Now consider the effect on a simple random layout of scatterers. By averaging over the phases alone, we will easily show that:

$$E(i) = \sum_{n=1}^{N_d} a_n E(\cos \phi_n) = 0 \quad E(q) = \sum_{n=1}^{N_d} a_n E(\sin \phi_n) = 0$$

$$E(i^2) = E(q^2) = \sum_{n=1}^{N_d} \frac{1}{2} a_n^2 = \frac{1}{2} R \quad E(iq) = 0$$

Parameter $R = \sum_{n=1}^{N_d} |e_n(u_n)|^2 |h(u_0 - u_n)|^2$ is proportional to the variance of z , as

$E(z) = 0$, and consequently $var(z) = E(i^2) + E(q^2)$. This corresponds to the reflective strength of the resolution cell. If an absolute calibration has been performed, R has the physical dimension of the radar cross-section (RCS), as shown in section 1.2.2.2. This naturally grows with the number of scatterers N_d and the mean scatterers $E(|e_n(u_n)|^2) = E(\sigma^0)$ RCS. When working with uncalibrated data, R is simply termed the radar reflectivity of a pixel. Also note that R can be interpreted as the incoherent image of the scatterer RCS ($\sigma^0 = |e|^2$) [OLI 91], whereas z is the coherent image of field e .

If N_d is sufficiently large, the real and imaginary parts i and q , sums of N_d independent random variables, are distributed according to a Gaussian distribution as a consequence of the central limit theorem. Finally, relation $E(iq) = 0$ reflects the lack of correlation between i and q . Thus, the two of them are independent because they are Gaussian with $p(i,q) = p_i(i|R)p_q(q|R)$ and:

$$p_i(i|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{i^2}{R}\right),$$

and

$$p_q(q|R) = \frac{1}{\sqrt{\pi R}} \exp\left(-\frac{q^2}{R}\right).$$

Since $E(z)$ mean is zero and curves $p(i,q) = \text{constant}$ are circles centered at the origin, it follows that z and $ze^{j\xi}$ have the same distribution regardless of phase shift ξ . Thus, speckle is a circular Gaussian process. The marginal distributions of the variables derived from z , particularly those of major radiometric variables such as the backscattered amplitude and intensity will now be developed. At pixel scale, fully developed speckle is characterized only by parameter R around which the backscattered intensity I shows large fluctuations due to slight motions of the scatterers. These fluctuations, when temporally observed, are called fading. We then have fully developed speckle, as data dispersion is maximal and as damaging as possible. Multi-look processing will tend to attenuate it, but will not remove it altogether.

In multi-dimensional cases, R is generalized by a covariance vector or matrix.

5.2.2.2. Distributions of amplitude A , phase ϕ and intensity I

By changing variables $i, q \rightarrow A = \sqrt{i^2 + q^2}, \phi = \tan^{-1}(q/i)$ we obtain $p(A, \phi|R) = p_A(A|R)p_\phi(\phi)$, where $p_\phi(\phi) = \frac{1}{2\pi}$ and $\phi \in [-\pi, \pi]$:

– full phase ϕ , which here is independent from amplitude and uniformly distributed, does not appear to carry any information; single channel data are therefore usually supplied as detected (generally amplitude A). In fact, it is not so at all and full complex signal z makes it possible to characterize the impulse response of the imaging system, i.e., the spatial speckle correlation (see section 5.3.2). We can use this information to obtain the optimal estimator of radar reflectivity (see section 6.2.2) that cannot be obtained from the detected data. By contrast, in the multi-dimensional case, the polarimetric and interferometric phase differences contain surface information as well;

– magnitude $A = |z|$ of the pixel value is distributed according to a Rayleigh distribution:

$$p_A(A|R) = \frac{2A}{R} \exp\left(-\frac{A^2}{R}\right) \quad A \geq 0 \quad [5.1]$$

and its mode is in $A = \sqrt{\frac{R}{2}}$, and the moments of order m equal $E(A^m) = R^{m/2} \Gamma\left(1 + \frac{m}{2}\right)$. In particular, the mean $E(A) = \sqrt{\frac{\pi R}{4}}$ and moment $E[A^2] = R$ enable us to deduce the amplitude coefficient of variation, which is a ratio of the standard deviation to the speckle mean:

$$\gamma_{SA} = \frac{E(A^2) - [E(A)]^2}{[E(A)]^2} = \sqrt{\frac{4}{\pi}} - 1 \approx \sqrt{0.2734} \approx 0.5227;$$

Parameter γ_S , which is sometimes called a contrast parameter, is a measure of speckle strength. In the case of non-homogenous surfaces, the radar reflectivity variations are integrated in the total coefficient of variation. It is used in both adaptive filtering (see Chapter 6) and texture analysis (see section 5.4). Figure 5.3 illustrates the Rayleigh distribution for two regions where reflectivity (radiometric) contrast $c = R_2/R_1$ is 3 dB. The figure shows very high radiometric confusion induced by speckle and a difficulty to discriminate between surfaces based on the value of a single pixel. To calculate the confusion rate, we use the assignment function $P(A < A_o | R) = 1 - \exp\left(-\frac{A_o^2}{R}\right)$;

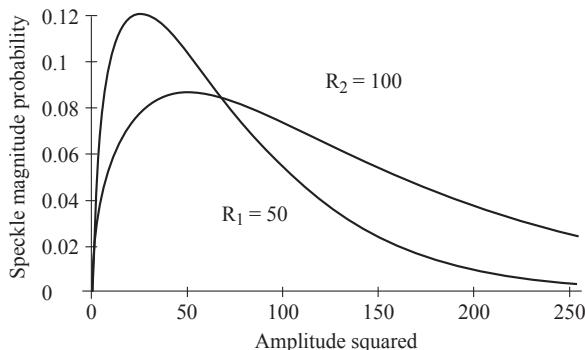


Figure 5.3. Amplitude distribution for two homogenous surfaces with radar reflectivities $R_1 = 50$ and $R_2 = 100$, respectively [BRU 96]

– the intensity $I = A^2$, which is proportional to the backscattered power and pixel luminance, is exponentially distributed (Laplace distribution) as:

$$p(I|R) = \frac{1}{R} \exp\left(-\frac{I}{R}\right) \quad I \geq 0. \quad [5.2]$$

This time, the most probable value is $I = 0$ (see Figure 5.4) and the moments are equal to $E(I^m) = R^m \Gamma(1+m)$. In particular, the mean value is $E(I) = R$, variance $\text{var}(I) = R^2$; and the single-look speckle coefficient of variation is $\gamma_{SI} = 1$. Intensity I is the efficient maximum likelihood (ML) estimator for R . Assignment function, $P(I < I_o | R) = 1 - \exp\left(-\frac{I_o}{R}\right)$ allows us to calculate the confidence rate we can assign to the single pixel value in order to estimate R . This will be illustrated in the general case of multi-look data later on (see section 5.3.4 and Figure 5.9). Note that the intensity, which is the sum of the squares of two independent Gaussian zero mean variables and identical variance $R/2$, can also be interpreted as a distributed variable, according to a chi-square distribution with 2 degrees of freedom, multiplied by $R/2$, wherefrom the sum of independent samples L mean distribution involved in the multi-look data should be easily deduced. Speckle, while still essentially Gaussian, is no longer fully developed in this case.

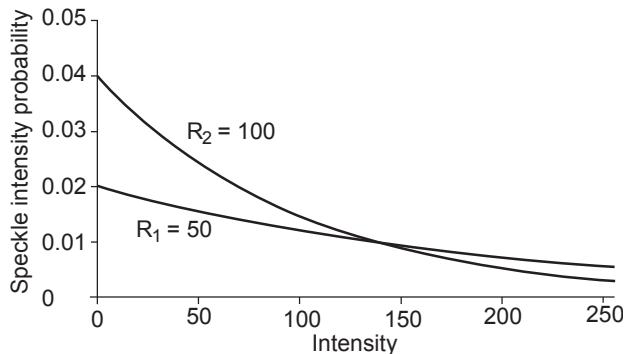


Figure 5.4. Distribution of the intensity I for two homogenous surfaces with radar reflectivities $R_1 = 50$ and $R_2 = 100$, respectively

5.2.2.3. Multiplicative and additive speckle models

According to the previous section, the standard deviation of data is proportional to expectation in both amplitude and intensity. This finding and conditional marginal distributions $p(.|R)$ explain why speckle is called multiplicative noise. This leads to:

$$\begin{aligned}
 z &= \sqrt{\frac{R}{2} S_z}, \\
 A = E(A)S_A &= \sqrt{\frac{R}{2} S_A}. \\
 I = E(I)S_I &= RS_I.
 \end{aligned} \tag{5.3}$$

where the speckle realizations in amplitude S_A and intensity S_I are distributed according to Rayleigh and Laplace distributions, respectively, with mean $E(S) = 1$.

Complex speckle S_z , defined as the normalized variance complex signal z , is distributed according to a complex circular Gaussian distribution with zero mean and parameter $R = 1$. It is as though R , which is the pixel's only accessible characteristic information, were modulated by interferences. Speckle reflects the way scatterers are organized inside every pixel involved. Such microscopic information cannot be obtained from a single image. By contrast, the comparison of speckle from two or more images is a source of additional information that we will exploit in interferometric and polarimetric studies.

Under certain conditions, the model can be spatially generalized to include both homogenous and non-homogenous surfaces as well as multi-dimensional data. Speckle can become additive via a logarithmic transformation of the amplitude or intensity image. However, for a multiplicative factor depending on the logarithm base b , density $\log_b I$ is $\log R + \log S_I$ and is distributed according to the Fisher-Tippett distribution (see equation [5.7]). The general expression for distribution is equation [5.7], section 5.3.4.3, applying to multi-look data. Logarithmic transformation has often been suggested for homomorphic speckle filtering by additive noise filters and other processing procedures such as boundary detection. In fact, this approach not only leads to a heavily biased estimation of radar reflectivity (by up to -2.507 dB), but also yields less than optimal results in terms of speckle variance reduction and, implicitly, in terms of performance rate, e.g. in detection (see Chapter 8). Its benefit, however, is found in the display of images where radiometric contrasts are enhanced due to constant variance. At high numbers of looks ($L > 10$), the distribution tends toward a Gaussian distribution. The real multi-variate Gaussian distribution can therefore be used to describe the joint distribution of several correlated channels. Although this distribution is only an approximation, it allows us to take into consideration the covariance of more than two channels in the classification problems using detected data – something that cannot be achieved with either the exponential and Rayleigh distributions or their multi-look equivalents, i.e. Gamma and generalized Gamma distributions. Once we use complex data, the multi-variate Gaussian distribution becomes accurate, which is one of the many benefits of SLC data, as will be seen in the following sections.

5.3. Complex circular multi-variate Gaussian model of vectorial or multi-look speckle

5.3.1. Joint distribution of single-look complex data

The complex circular Gaussian model can be generalized to N measures that may be either $N = L$ sub-looks (see section 5.3.4), N frequency or polarization channels, or a set of N neighboring pixels of the same image. In this case, z turns into a complex vector \mathbf{Z} where N components are the z_n measures in the different channels. If the Goodman hypotheses are verified for each of these channels, and the imaginary and real parts of z_n verify the two following conditions:

$$\begin{aligned} E(Re(z_m)Re(z_n)) &= E(Im(z_m)Im(z_n)) \\ E(Re(z_m)Im(z_n)) &= -E(Im(z_m)Re(z_n)) \end{aligned}$$

the distribution of \mathbf{Z} is [GOO 63]:

$$p_z(\mathbf{Z}|\mathbf{C}_z) = \frac{1}{\pi^N \det(\mathbf{C}_z)} \exp(-{}^t \mathbf{Z}^* \mathbf{C}_z^{-1} \mathbf{Z}) \quad [5.4]$$

where ${}^t \mathbf{Z}$ is the transposed vector of \mathbf{Z} , and $\mathbf{C}_z = E(\mathbf{Z} {}^t \mathbf{Z}^*)$ is the covariance matrix of \mathbf{Z} . The Hermitian matrix \mathbf{C}_z generalizes the radar reflectivity concept, since its diagonal elements, i.e., z_n variances, correspond to N radar reflectivities R_1, R_2 , etc., R_N . However, the non-diagonal elements $C_{nm} = cov(z_n, z_m)$, which are covariances between the components, are proportional to the complex correlation coefficient (also referred to as the degree of coherence) between z_n and z_m :

$$\rho_{nm} = \frac{C_{nm}}{\sqrt{C_{nn}C_{mm}}} = \frac{E(z_n z_m^*)}{\sqrt{E(|z_n|^2)E(|z_m|^2)}} = D_{nm} \exp(j\beta_{nm})$$

The complex correlation module D is simply called coherence, while the phase β is the effective phase shift between \mathbf{Z} components. It is this particular parameter that is sought in radar interferometry, as the phase difference becomes proportional to the path difference between the radar and the resolution cell in both acquisitions. Yet, due to speckle, the phase difference between measures z_n and z_m is not equal to β . As a result, estimators, generally the ML estimator, have to be used. The

distribution of the observed phase difference and that of the coherence ML estimator are given as multi-look data are discussed further below.

5.3.2. Spatial correlation and spectral properties of the single-look Gaussian speckle

Speckle, as previously discussed, may be considered as the response to a complex linear filter $h(u)$ of the following signal:

$$\epsilon(u) = \sum_n e_n(u_n) \cdot \delta(u - u_n) = \sum_n a_n e^{j\varphi_n} \cdot \delta(u - u_n).$$

In these conditions, signal $\epsilon(u)$ can be modeled as a white complex circular Gaussian process of autocorrelation function:

$$C_\epsilon(u_1 - u_2) = E\left(\epsilon(u_1)\epsilon(u_2)^*\right) = \sqrt{\sigma^o(u_1)\sigma^o(u_2)} \cdot \delta(u_1 - u_2),$$

where $\sigma^o(u)$ is the density of radar reflectivity per unit area [MAD 87]. For a homogenous (broadly stationary) surface, $\sigma^o(u)$ is constant. It also represents the power spectral density of the signal. By introducing here the effect of range D in phase φ_n , $z(u_o)$ can be written:

$$z(u_o) = \sum_{n=1}^{N_d} e_n(u_n) \cdot h(u_0 - u_n) = \sum_{n=1}^{N_d} a_n e^{j\varphi_n} \cdot h(u_0 - u_n).$$

– mean intensity:

$$E(R) = E(I(u_o)) = E\left(z(u_o)z^*(u_o)\right)$$

is equal to:

$$E(R) = E\left(a_n^2\right) \sum_{n=1}^{N_d} |h(u_n)|^2 = \sigma^o \sum_{n=1}^{N_d} |h(u_n)|^2 ;$$

– the complex field autocorrelation function is:

$$C_z(\Delta u) = E\left(z(v)z^*(v + \Delta u)\right) = \sigma^o \sum_{n=1}^{N_d} h^*(u_n)h(u_n - \Delta u);$$

and hence the spatial autocorrelation complex coefficient of z , which is the same as that of complex speckle S , is:

$$\rho_S(\Delta u) = \frac{C_z(\Delta u)}{E\left(z(u_o)z^*(u_o)\right)} = \frac{\sum_{n=1}^{N_d} h^*(u_n)h(u_n - \Delta u)}{\sum_{n=1}^{N_d} |h(u_n)|^2}$$

Thus, it only depends on the complex impulse response of the sensor. The same property holds even where the surface is no longer stationary [MAD 87];

– the autocovariance function of intensity is:

$$C_I(\Delta u) = E(I(v)I^*(\Delta u + v)) = (\sigma^o)^2 \left[\sum_{n=1}^{N_d} |h(u_n)|^2 + \left| \sum_{n=1}^{N_d} h^*(u_n)h(u_n - \Delta u) \right|^2 \right]$$

Similarly, intensity spatial correlation coefficients $\rho_I(\Delta u)$ depend on nothing but the sensor's complex impulse response. Furthermore, they are always positive (according to the Siegert relation):

$$\rho_I(\Delta u) = |\rho_S(\Delta u)|^2.$$

This property no longer holds unless the surface is homogenous [BLA 93]. Note that we can find the same properties by modeling the signal $c(u)$ at the filter $h(u)$'s entry as a white continuous random Gaussian process with zero mean and a power spectral density σ^o . The output $z(u)$ is then a narrow band Gaussian process. The correlation function is expressed as:

$$C_z(\Delta u) = \sigma^o \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h^*(u) h(u - \Delta u) du.$$

Also note that if $h(x,y)$ is a separable function (see section 2.2.4), we can split the two-dimensional spatial correlation coefficient of complex speckle into two one-dimensional coefficients $\rho_S(\Delta x, \Delta y) = \rho_S(\Delta x)\rho_S(\Delta y)$, where correlation in each direction is:

$$\rho_S(\Delta u) = \frac{\int_{-\infty}^{+\infty} h^*(u) h(u - \Delta u) du}{\int_{-\infty}^{+\infty} |h(u)|^2 du};$$

– the complex signal power spectral density is given by the relation:

$$S_Z(f) = \sigma^o |H(f)|^2$$

where $H(f)$ is the Fourier transform of complex impulse response $h(u)$. Since $H(f)$ is bound, $h(u)$ is theoretically infinite, or at least spreads over a large area, with strong overlap of neighboring resolution cells. Therefore, in SAR images, spatial correlation is generally not zero;

– spatial sampling and correlation

Let us consider a radar image sampled at a step Δ_u along considered direction u . In the case of a limited complex spectral image, which is here limited by filter $H(f)$, the Δ_u sampling (pixel dimension) can be conducted to the Nyquist-Shannon theorem limit. For instance, for a gate filter $H(f)$ with a width of $F_e = 1/\Delta_u$, we have relation $\delta_u = 0.885\Delta_u \approx \Delta_u$ between the spatial resolution at 3 dB and sampling step Δ_u . For detected images, Δ_u has to be divided by 2 because the spectral width is doubled. In practice, this half-step is often used even for SLC images. As a result, distances between pixels are multiples of $\frac{1}{2}\delta_u$, i.e. $\Delta u = n\Delta_u \approx n\frac{1}{2}\delta_u$. Figure 5.5 illustrates spatial correlations in a complex ERS image up to $n = 15$ in both directions.

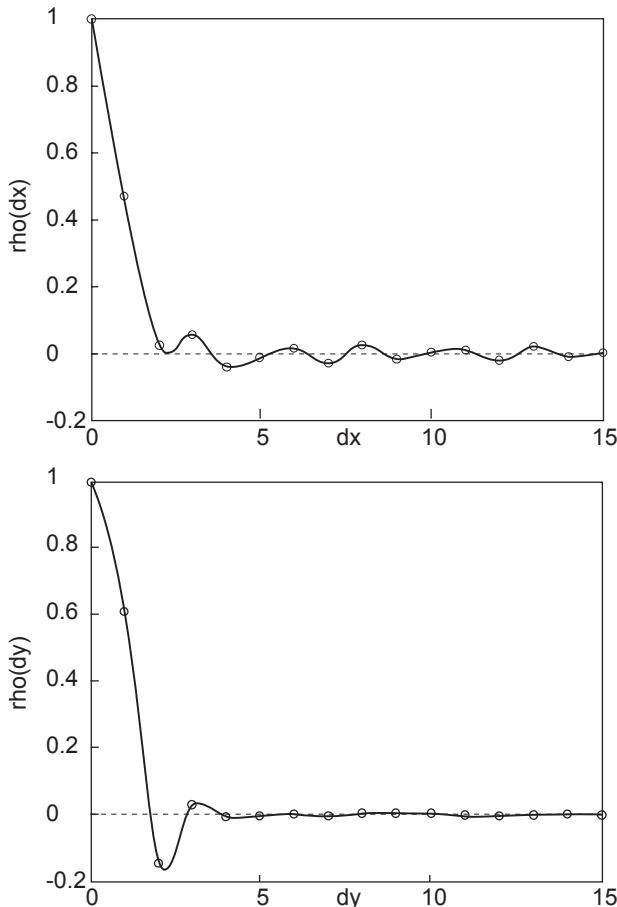


Figure 5.5. ERS speckle spatial correlations over radial range (top) and azimuth (bottom)

These practically real correlations correspond to the spectra shown in Figure 5.6, which only cover 80% of the total bandwidth. The first spectrum reflects weighting for antenna pattern in azimuth and a Hamming filter. The latter corresponds to the range chirp and makes it possible to achieve good spatial resolution after impulse compression (see section 2.2.1). The reduced occupation rate of the spectra (here about 80%) and weighting for the antenna pattern and filters account for the high value of the correlation of 2 successive pixels. The correlation could be neglected beyond $n = 3$ pixels. In fact, for this correlation to be zero, aside from case $\Delta_u = 0$, the speckle spectrum would have to be white, i.e. take 100% of useful band. Indeed, if $|H(f)|^2$ is

both the gate function of width $Fe = 1/\Delta_u, C_z(\Delta_u)$ and inverse Fourier transform of $|H(f)|^2$, the spatial correlation is simply $\rho_S(n\Delta u) = \frac{\sin n\pi}{n\pi}$.

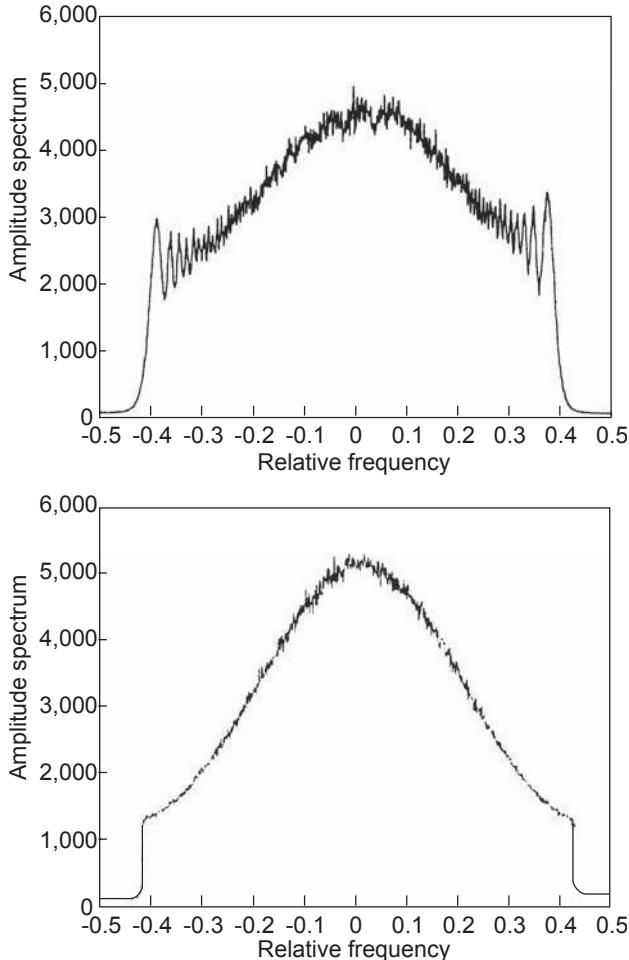


Figure 5.6. ERS speckle spectrum corresponding to radial (top) and azimuth (bottom) range correlations in Figure 5.5

To remove spatial correlation, we simply reverse the zero-padding procedure used in interpolation. To this end, image resampling and spectral whitening are carried out. This involves a fast Fourier transform, followed by suppressing the zero part of the spectrum, equalizing the spectrum amplitude and inverting the FFT. The

image size is thus reduced, while the pixel size grows proportionally without any loss of information. By contrast, point targets (or staircase-like transitions in radar reflectivity R) lead to a very high amount of side lobes in the response. Removing the spatial correlation of speckle improves radar reflectivity estimates, which are simplified and become more accurate as a result. The R estimator, using the simple arithmetic mean of intensities I based on detected data, becomes effective at this point. The same may be true for other radar parameters as well, given that estimators are generally established only for independent samples and independence is equivalent to decorrelation for two variables with the Gaussian, Gamma and generalized Gamma distributions characteristic of speckle in different data formats. If no spectral whitening is applied, the spatial properties of speckle can still be used, but only based on SLC complex data. This is illustrated in the optimal estimation of radar reflectivity (see [FJO 99b] and section 6.2.2), which, for example, leads to a substantial improvement in detection results (see sections 8.1 and 8.2). On the other hand, spectral whitening enhances the effects of thermal noise. The additive thermal noise, or white Gaussian noise, shares the statistical properties of speckle, except that its spectral properties are different, since it does not go through the entire radar system, particularly in the azimuth direction where it avoids antenna pattern weighting. Except in the case of gate function filtering and weighting, it can modify the spatial correlations of speckle, which theoretically depend on the radar system alone, as well as correlations of the applied filters and sampling step. In all these situations, the complex spatial correlations can be estimated over broad image areas in order to obtain accurate unbiased estimates. The ML estimator of a complex correlation coefficient for independent samples is given later in this chapter (section 5.3.5.1). Unlike in the case of detected images, the surfaces do not need to be homogenous [MAD 87].

In other effects, thermal noise skews radar reflectivity towards higher values and hampers the precision of radar reflectivity estimates. If its backscattering coefficient equivalent σ^0 is noise equivalent σ_{ne}^0 , the radar backscattering coefficient σ^0 of a surface will be biased by $10 \log(1 + \sigma_{ne}^0/\sigma^0)$. On the ERS images, $\sigma_{ne}^0(\text{dB}) \approx -23 \text{ dB}$ and can often be ignored up to a signal-to-noise ratio σ_{ne}^0/σ^0 higher than 10, the bias for which is lower than 0.5 dB.

5.3.3. Joint distribution of single-look detected data

Unlike the complex case, the distribution of a vector $\mathbf{Y}_A = (A_1, A_2, \text{etc.}, A_N)$ (or $\mathbf{Y}_I = (I_1, I_2, \text{etc.}, I_N)$, respectively), where components are realizations of N correlated random variables distributed according to a Rayleigh distribution (or an exponential distribution, respectively) has no analytical expression upwards of $N = 3$ in general [MIL 74]. An analytical expression could be found up to $N = 5$ pixels, but only under restrictive assumptions as to the shape of the covariance matrix which cannot

accurately represent the spatial properties of the SAR image speckle. This expression has nonetheless been used somewhat successfully in the Bayesian segmentation of single-look radar images [KEL 99, QUE 93a]. A segmentation of SLC images based on a more appropriate modeling is given in [RIG 92].

However, in the case $N = 2$, joint distribution can be obtained regardless of the situation and only based on coherence D between the channels and the two reflectivities R . A more general expression is given in the case of multi-look data (see section 5.3.4).

5.3.4. Distributions of scalar multi-look data

The reduction of speckle, which is then called “non-fully developed”, by multi-look processes is obtained by dividing the Doppler spectrum into L sub-looks. The looks generally overlap and the L sub-looks z_n will then be correlated. However, as an effect of the weighting filters and antenna pattern, the L realizations of the sub-look speckle become virtually decorrelated, hence independent, when looks overlap too much as they do in Figure 5.7. Between looks n and m , as obtained with filters $H_i(f)$ corresponding to the impulse response $h_i(u)$, speckle covariance is given by the interference formula:

$$C_{nm}(u) = \sigma^0 h_n(u) * h_m^*(-u).$$

The impulse responses are functions of type:

$$h_n(u) = A(u) \exp\left(-j \frac{2k}{D_0} l_n u\right) \exp\left(-j \frac{k}{D_0} u^2\right),$$

where l_n / \tilde{L} is the relative shift between the center of the considered synthetic antenna and the center of the nominal synthetic antenna corresponding to full spatial resolution with a length of $\tilde{L} \approx \beta D_0$. β is the aperture of the real antenna pattern; D_0 is the radial distance from the center of the pixel and $k = \frac{2\pi}{\lambda}$. Every sub-look is obtained from a synthetic antenna, with a length $l = \eta \tilde{L}$ that is a fraction η of the nominal synthetic antenna \tilde{L} . $A(u)$ is the amplitude response. If this is an even function, the coherence ρ_{nm} is real. For instance, if filter $H_n(f) = I_{\delta_f}(f)$ is a gate function with a

width of $\delta_f = \frac{2l}{\lambda D_0}$ then $A(u) = \begin{cases} \sin \frac{kl}{D_0} u & \\ \frac{kl}{D_0} u & \end{cases}$ with a spatial resolution within -3 dB of $\delta u \approx 0.885 \frac{\pi D_0}{kl} \approx \frac{\pi D_0}{kl}$. The correlation between two looks then depends linearly on their overlap percentage $\tau = \frac{|l_n - l_m|}{l}$ provided they are separated by a $|l_n - l_m|$ smaller gap than l . The correlation value in this case is $\rho(\tau) = 1 - \tau$. If extraction is performed through Gaussian filters with a width of l at $\frac{1}{e}$, the correlation between looks is $\rho(\tau) = \exp\left(-\frac{\tau^2}{2}\right)$. Although Gaussian modeling is less than rigorous because $H_n(f)$ filters are bandlimited, it becomes possible to obtain analytical results facilitating the analysis and understanding of many processes and phenomena specific to SAR imagery [RAN 80, OUC 87, POS 93]. In practice, the filters are often Taylor (Hamming) devices that help reduce secondary lobes considerably. The correlation between looks for a Hamming filter is:

$$\rho(\tau) = \frac{2a^2(1-\tau) + \frac{2a(1-a)}{\pi} \sin(2\pi\tau) + (1-a)^2 \left[(1-\tau) \cos(2\pi\tau) - \frac{1}{2\pi} \sin(2\pi\tau) \right]}{1 - 2a + 3a^2}$$

where $a \in [0; 0.5]$.

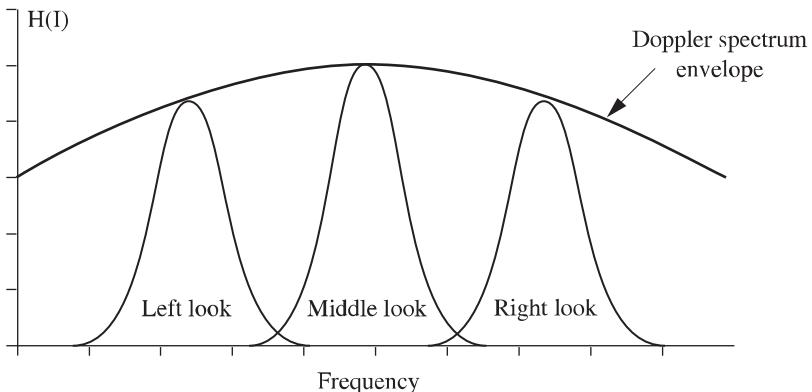


Figure 5.7. Example of extraction of 3 virtually uncorrelated looks by filters $H_n(f)$ with the same bandwidth

5.3.4.1. Gamma and generalized Gamma distributions

The multi-look image with an intensity $I = \frac{1}{L} \sum_{k=1}^L I_k$ is distributed according to a

Gamma distribution if energy in the sub-looks is constant (we can always equalize it) and correlations between looks are zero. The Gamma distribution is the distribution of a sum of independent L variables distributed according to the distribution of χ^2 with two degrees of freedom. Therefore, it is the distribution of a random variable of χ^2 with $2L$ degrees of freedom, multiplied by $R/2L$. It has two parameters: a parameter of scale R and a parameter of shape L :

$$p_I(I|R) = \left(\frac{L}{R}\right)^L \frac{1}{\Gamma(L)} \exp\left(-\frac{LI}{R}\right) I^{L-1} \quad I \geq 0. \quad [5.5]$$

The moments are given by relation:

$$E(I^m) = R^m \frac{\Gamma(L+m)}{L^m \Gamma(L)}.$$

It follows that $E(I) = R$ and $\text{var}(I) = \frac{1}{L} R^2$ meaning that the multi-look image does not change radar reflectivities relative to the single-look image and the coefficient of variation is $\gamma_{SI} = \frac{1}{\sqrt{L}}$. Speckle variance is reduced by a factor L , but this improved radiometric resolution is obtained at the expense of spatial resolution. Figure 5.8 shows that dispersion around R mean does not decrease significantly unless we have a high number of looks L .

More specifically, we can determine the probability that one pixel belonging to a distribution according to the Gamma distribution area differs from mean value $R = E(I)$ by a relative deviation $\eta \leq 1$, i.e., the probability that $I \notin [R(1 - \eta), R(1 + \eta)]$. This probability of error, defined by relation:

$$P_E = \int_0^{R(1-\eta)} p_I(I|R) dI + \int_{R(1+\eta)}^{+\infty} p_I(I|R) dI$$

is given by [BRU 96]:

$$P_E = \frac{1}{\Gamma(L)} \left[\Gamma(L) + \Gamma(L, L(1+\eta)) - \Gamma(L, L(1-\eta)) \right]$$

where $\Gamma(\alpha, z)$ is the incomplete Gamma function. Figure 5.9 shows that for a deviation η of $\pm 10\%$ (or about ± 0.5 dB), we cannot obtain a probability of error that is reasonably low unless the number of looks L is in the hundreds!

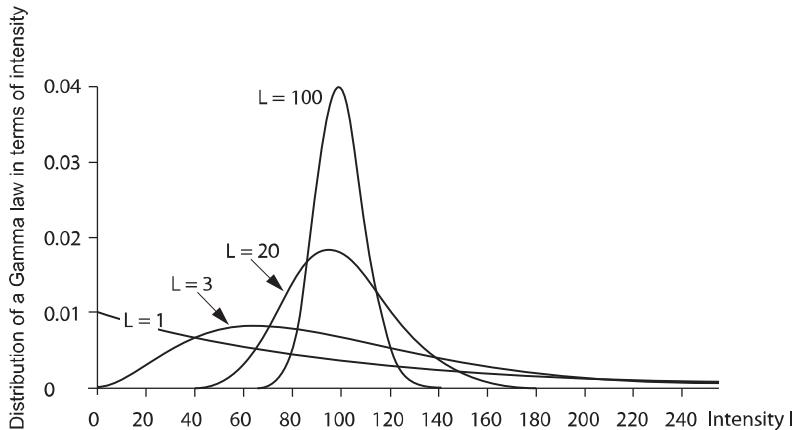


Figure 5.8. Intensity speckle distribution for an L -look image
in reflectivity area $R=E(I)=100$

The distribution law in terms of amplitude $A = \sqrt{I}$ is given by the generalized Gamma distribution:

$$p_A(A|R) = 2 \left(\frac{L}{R} \right)^L \frac{1}{\Gamma(L)} \cdot \exp \left(-\frac{LA^2}{R} \right) A^{(2L-1)}, \quad A \geq 0,$$

with the mean $E(A) = \sqrt{\frac{\pi R}{4}}$, moments $E(A^m) = R^{\frac{m}{2}} \frac{\Gamma(L+\frac{m}{2})}{L^{\frac{m}{2}} \Gamma(L)}$ and an amplitude coefficient of variation $\gamma_{SA} = \sqrt{\frac{L\Gamma^2(L)}{\Gamma^2(L+\frac{1}{2})}} - 1$.

For $L = 3$, coefficient of variation γ_{SA} amounts to approximately 0.2941. Changing the variable will obviously not change the probabilities of error calculated on intensity.

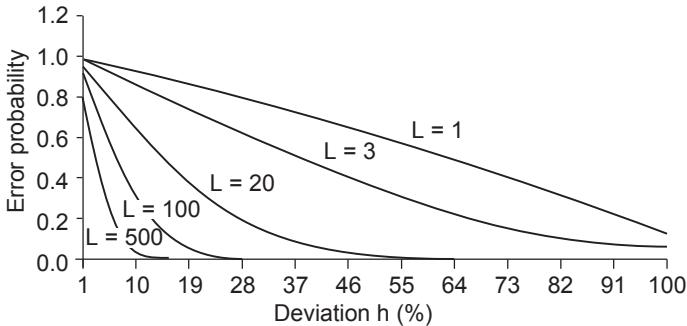


Figure 5.9. P_E estimation error of reflectivity R based on a sample of intensity I in an L -look image according to relative deviation from actual reflectivity R [BRU 96]

The probability density of speckle in amplitude multi-look images performed directly in amplitude $A = \frac{1}{L} \sum_{k=1}^L A_k$ is the successive convolution of L Rayleigh distributions. There is no analytical expression for it, but we know its moments:

$$E(A^m) = R^{m/2} \frac{1}{L^m} \Gamma\left(1 + \frac{m}{2}\right),$$

and therefore

$$E(A) = \sqrt{\frac{\pi R}{4}},$$

$$\gamma_{SA} = \frac{1}{\sqrt{L}} \sqrt{\frac{4}{\pi}} - 1 \approx \frac{0.5227}{\sqrt{L}}.$$

For $L = 3$, $\gamma_{SA} \approx 0.3018$, which is rather close to the characteristics of the generalized Gamma distribution. Once $L > 3$, this distribution can be used in practice, whatever the type of amplitude multi-look image we are dealing with.

5.3.4.2. Goodman distribution

If the multi-look sum is carried out with correlated looks, we obtain an L' -look image where $L' \leq L$. Moreover, intensity is no longer distributed according to a Gamma distribution, but rather according to a different distribution that we will call the Goodman distribution. The intensity distribution of I is expressed in terms of L real non-negative eigenvalues λ_k of the covariance matrix [GOO 75]. This distribution may also apply to the mean spatial intensity of $L = N$ correlated pixels of a single-look image and a homogenous surface. If reflectivity R is identical for all L looks or pixels, the Goodman distribution is expressed as:

$$p_I(I|R) = \sum_{k=1}^L a_k e^{-\frac{I}{R\lambda'_k}}, \quad a_k = \frac{L}{R\lambda'_k \prod_{j \neq k}^L \left(1 - \frac{\lambda'_j}{\lambda'_k}\right)}.$$

In these conditions, $\lambda'_k = \frac{\lambda_k}{R}$ are L real eigenvalues of the correlation matrix, whose non-diagonal elements ρ_{nm} are the coherences between L components. Covariance matrix \mathbf{C}_z is expressed as:

$$\mathbf{C}_z = R \begin{bmatrix} 1 & \rho_{12} & \dots & \rho_{1L} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \rho_{1L}^* & \dots & \dots & 1 \end{bmatrix}$$

For $L = 2$ looks or $N = 2$ pixels, the mean of 2 intensities is written according to the coherence degree $D = |\rho_{12}|$ and R :

$$p(I|R,D) = \frac{2}{RD} \exp\left[-\frac{2I}{R(1-D^2)}\right] \sinh\left[-\frac{2ID}{R(1-D^2)}\right]$$

where \sinh is the hyperbolic sine function. The Goodman distribution, not easy to use for $L > 2$, is almost always approached in practice by a Gamma distribution using L' rather than L as form parameter. The effective (or equivalent) number L' of independent looks, which is a real number rather than an integer like L , is defined

here from relation $\gamma_{SI} = \frac{1}{\sqrt{L'}}$ which is identical to the one between the coefficient of variation and the number of L looks in the case of decorrelated looks. The equivalent L' number of looks is:

$$L' = \frac{L}{1 + \frac{2}{L} \sum_{n=1}^{L-1} \sum_{m=n+1}^L \rho_{Inm}}. \quad [5.6]$$

Since intensity correlation coefficients ρ_{Inm} , given by $\rho_{Inm} = |\rho_{Inm}|^2$, are always positive, the equivalent number of looks is always lower than L , except where all the correlations are zero and where $L' = L$. Where overlap $\tau = \frac{|l_n - l_m|}{l}$ is limited to 50%, which is common practice, the correlation is never non-zero except for two adjacent sub-looks. For example, for $L = 3$ and $\rho_l = D^2 = 0.25$ (maximum correlation with 50% overlap and gate filters), $L' = 2.25$ and, finally, if $L = 2$, then $L' = 1.6$. If ρ_l values are unknown, L' has to be estimated using either the empirical coefficient of variation, or the ML estimator of a Gamma distribution form parameter, but only over extensive homogenous areas. Figure 5.10 shows a comparison between the Goodman distribution, L' parameter Gamma distribution, and L parameter Gamma distribution for $L = 2$ looks and a degree of coherence of looks $D = 0.5$. The approximation provided in this case by L' parameter Gamma distribution is merely satisfactory. It may improve, though, as L grows or as correlations decrease. Moreover, as L' becomes higher than 5, it also helps approximate the distribution of ratio r of intensity means [BRU 98]. This is a variable that is widely used for target and boundary detection (see section 8.1.2), for the polarization ratio in polarimetry and for studying surface changes in the multi-temporal approach. We will use hereinafter the equivalent number approximation of independent looks (pixels); the number of looks, whether integer or real L' , will always be denoted by L .

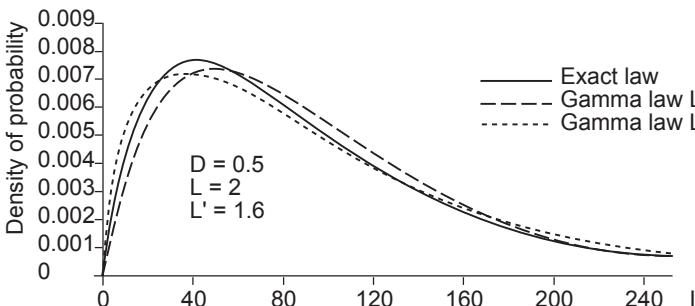


Figure 5.10. Comparison of multi-look distributions for case $L = 2$ and $D = 0.5$

5.3.4.3. Multi-look logarithmic data

In the case of multi-look data, $D_L = \log I$ is distributed according to the Fisher-Tippett distribution:

$$p_D(D_L|R,L) = \frac{L}{\Gamma(L)} \exp[-L(D_r - D_L)] \cdot \exp[-L \exp(-(D_r - D_L))], \quad [5.7]$$

where $D_r = \ln R$.

The mean $E(D_L) = D_r - \ln L + \Gamma'(L)/\Gamma(L)$ shows that a negative bias has been introduced on R . This bias is maximum for $L = 1$ look, with $E(D_L) = D_r - \Upsilon$ and Euler's constant $\Upsilon \approx 0.577$, i.e. a bias of -2.507 dB. Variance is:

$$\text{var}(D_L) = \frac{\pi^2}{6} - \left[\Upsilon + \frac{\Gamma'(L)}{\Gamma(L)} \right]^2 + 2 \sum_{k=1}^{L-2} \left(\frac{1}{L-k} \sum_{j=1}^{L-k-1} \frac{1}{j} \right).$$

which is also maximal for $L = 1$ and where $\text{var}(D_L) = \frac{\pi^2}{6}$, i.e., a standard deviation of 5.57 dB. Figure 5.11 shows the distribution of D_L for different numbers of looks. The Gaussian distribution approximation only works satisfactorily from ten looks upwards.

If a homomorphic transformation is used, then finding the arithmetic mean $\ln I$ over N samples comes down to calculating random variable $I_G = \exp\left(\frac{1}{N} \sum_{n=1}^N \ln I_n\right)$,

which is also the geometric mean of samples $I_G = \prod_{n=1}^N I_n^{1/N}$. The distribution of this variable is unknown, but we can calculate its moments if N samples are independent. It should be noted that the mean differs from real R reflectivity. The bias is provided by the multiplicative factor $B_G < 1$:

$$E(I_G) = R \cdot B_G = R \cdot \frac{1}{L} \frac{\Gamma^N\left(L + \frac{1}{N}\right)}{\Gamma^N(L)}$$

its extreme value is 0.561 (or -2.507 dB) for $L = 1$ and infinite N (see Figure 5.12).

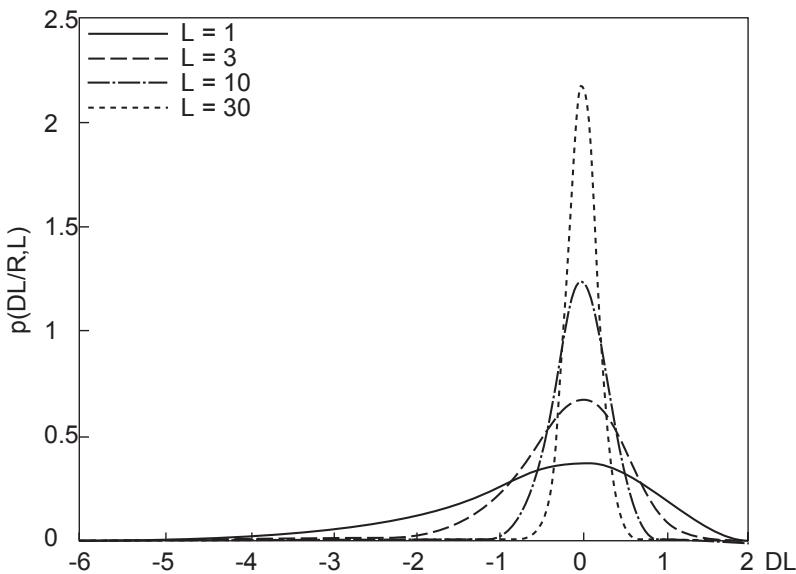


Figure 5.11. Speckle distribution following logarithmic transformation ($D_r = 0$)

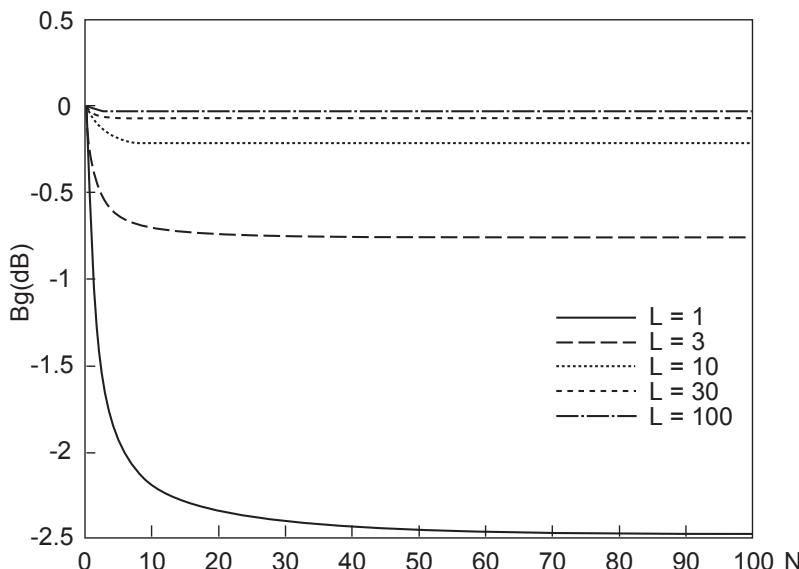


Figure 5.12. Bias (in log) of geometric mean samples

It only becomes zero if L is infinite or $N = 1$. For a fixed number of looks L , it grows with the number of independent N samples. To obtain an unbiased quantity, we will have to divide I_G by B_G , which poses no difficulty in practice since L and N are fixed. However, the variance of this corrected variable $I'_G = \frac{L\Gamma^N(L)}{\Gamma^N(L+\frac{1}{N})} I_G$ is always higher than that of the optimal radiometry R ML estimator, which is the mere arithmetic mean of intensities, and where variance $\frac{R^2}{NL}$ achieves the Cramer-Rao lower bound. The standard deviation increase is given by the multiplicative factor $f_{G'} > 1$:

$$f_{G'}' = NL \left[\frac{\Gamma^N\left(L + \frac{2}{N}\right) \Gamma^N(L)}{\Gamma^{2N}\left(L + \frac{1}{N}\right)} - 1 \right].$$

Number $f_{G'}$ illustrated in Figure 5.13 can be interpreted as shown by a relative increase of the independent sample number which is necessary to reach the same variance as that obtained using the arithmetic mean of the intensities. For example, for $L = 1$ look and a high N (worst case scenario), we have $f_{G'}' \approx \frac{\pi^2}{6} = 1.645$. Therefore, we will need to add another 65% of samples if we use the geometric mean, while if we use the arithmetic mean of amplitudes $f_{A'} = 4(4/\pi - 1) \approx 1.093$, the sample number will only have to be increased by about 9% [ZRN 75]. These results hold even in the case of correlated samples [FJO 99a]. This explains why, to estimate a radar reflectivity with independent samples, particularly if the number of looks L is small, we had better use the mean intensity, or even the mean amplitude if necessary, rather than the logarithmic (or dB) mean. By contrast, if the samples are correlated, SLC complex data should be preferred if available (see section 6.2.2). This logarithmic image drawback is also found in image processing problems such as boundary detection [FJO 99a], though not in multi-look data classification problems where it allows us to use the multi-variate Gaussian distribution.

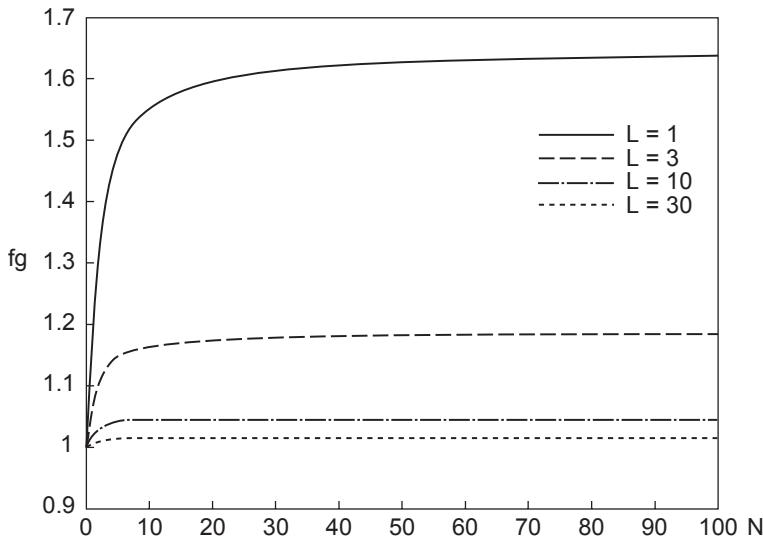


Figure 5.13. Relative increase in sample numbers required to estimate R reflectivity following a log mean

5.3.4.4. Spectral and spatial properties of multi-look speckle

As previously described, spatial correlations of intensity speckle in every look are simply equal to the square module of the corresponding complex speckle coherences. If L looks are statistically identical and obtained via the same filter $H_L(f)$, we can easily demonstrate that the spatial correlation coefficients of multi-look intensity image speckle are always given by $\rho_{SI}(\Delta u) = |\rho_z(\Delta u)|^2$. Likewise, the power spectral density of speckle S_I is easily derived from expression $I = zz^*$ and is thus proportional to $|H_L(f)^* H_L^*(-f)|^2$. In the ideal case of a gate function, spectral density is thus a triangle function, with a width that is twice that of function $H_L(f)$. Detected images have to be sampled at half the step used in SLC images with the same spatial resolution.

5.3.5. Multi-look vectorial distributions

Now consider L samples \mathbf{Z}_k of measurement vector \mathbf{Z} with dimension N , identically distributed according to a multi-variate complex circular Gaussian distribution with a covariance matrix \mathbf{C}_Z . The empirical covariance matrix $\mathbf{\Sigma}_Z$ of

dimension $N \times N$ (the form to which multi-look data can always be converted even in polarimetry) is defined by:

$$\Sigma_{\mathbf{Z}} = \frac{1}{L} \sum_{k=1}^L \mathbf{Z}_k {}^t \mathbf{Z}_k^*.$$

If the samples are independent, $\Sigma_{\mathbf{Z}}$ is the ML estimator of $\mathbf{C}_{\mathbf{Z}}$ and is distributed according to the complex Wishart distribution [GOO 63]:

$$p_{\Sigma} = (\Sigma_{\mathbf{Z}} | \mathbf{C}_{\mathbf{Z}}) = \frac{L^{LN} |\Sigma_{\mathbf{Z}}| L^{-N} \exp(-L \cdot \text{Tr}(\mathbf{C}_z^{-1} \Sigma_z))}{\pi^{N(N-1)/2} \Gamma(L) \Gamma(L-1) \text{ etc. } \Gamma(L-N+1) |\mathbf{C}_z|^L}. \quad [5.8]$$

The diagonal elements I_n of $\Sigma_{\mathbf{Z}}$, homogenous to intensities, are distributed according to a Gamma distribution with parameters R_n and L . A non-diagonal element is called a Hermitian product. The complex Wishart distribution is a generalization of the Gamma distribution for complex multi-look multi-dimensional data. Further parameters are often extracted from $\Sigma_{\mathbf{Z}}$. Their definitions and distributions are given below.

5.3.5.1. Distribution of variables derived from $\Sigma_{\mathbf{Z}}$ and the Wishart distribution

The most commonly used variables, including phase difference, empirical coherence, channel ratio and Hermitian product, can be defined in the particular case of $N=2$ channels. The joint distribution of elements of:

$$\Sigma_{\mathbf{Z}} = \begin{pmatrix} I_1 & I_{12} e^{j\varphi} \\ I_{12} e^{-j\varphi} & I_2 \end{pmatrix}$$

is expressed in terms of reflectivities R_1 and R_2 , the degree of coherence $\rho = D e^{j\beta}$ between channels and the number of looks L as:

$$p(I_1, I_2, I_{12}, \varphi | R_1, R_2, D, \beta, L) = \frac{L^{2L} (I_1 I_2 - I_{12}^2)^{L-2}}{\pi (1-D^2)^L R_1^L R_2^L \Gamma(L) \Gamma(L-1)} \cdot \exp \left[-\frac{L}{1-D^2} \left(\frac{I_1}{R_1} + \frac{I_2}{R_2} - \frac{2I_{12}D(\cos(\varphi - \beta))}{\sqrt{R_1 R_2}} \right) \right] \quad [5.9]$$

We can deduce the distributions of the Hermitian product of module I_{12} , empirical coherence $\hat{\rho} = \frac{I_{12}e^{j\varphi}}{\sqrt{I_1 I_2}} = de^{j\varphi}$ of its module d (degree of empirical coherence) and phase φ (empirical phase difference), by only showing the parameters on which they depend [LEE 94a, TOU 96, BRU 96, TOU 99]:

– distribution of module I_{12} of the Hermitian product:

$$p(I_{12}) = \frac{4L^{L+1} I_{12}^L}{(1-D^2)\Gamma(L)(\sqrt{R_1 R_2})^{L+1}} I_0\left(\frac{2LDI_{12}}{(1-D^2)\sqrt{R_1 R_2}}\right) \cdot K_{L-1}\left(\frac{2LI_{12}}{(1-D^2)\sqrt{R_1 R_2}}\right) \quad [5.10]$$

where $I_0(x)$ and $K_{L-1}(x)$ are modified Bessel functions of the first and second types;

– polar distribution of the empirical coherence:

$$p(d, \varphi | D, \beta, L) = \frac{(1-D^2)^L \Gamma^2(2L)}{2^{4L-2} \Gamma(2L + \frac{1}{2}) \Gamma(L) \Gamma(L-1) \Gamma(\frac{1}{2})} d(1-d^2)^{L-2} \cdot {}_2F_1\left(2L, 2L; 2L + \frac{1}{2}; \frac{1}{2}(1+dD \cos(\varphi - \beta))\right) \quad [5.11]$$

– empirical phase difference distribution:

$$p(\varphi | D, \beta, L) = \frac{(1-D^2)^L}{2\pi} \left({}_2F_1\left(1, L; \frac{1}{2}; D^2 \cos^2(\varphi - \beta)\right) + D \cos(\varphi - \beta) {}_1F_0\left(L + \frac{1}{2}; -; D^2 \cos^2(\varphi - \beta)\right) \frac{\Gamma(\frac{1}{2}) \Gamma(L + \frac{1}{2})}{\Gamma(L)} \right) \quad [5.12]$$

– distribution of the empirical degree of coherence d :

$$p(d|D, L) = 2(L-1)(1-D^2)^L d(1-d^2)^{L-2} {}_2F_1(L, L; 1; d^2 D^2) \quad [5.13]$$

for $L > 1$, where ${}_2F_1(\cdot, \cdot; \cdot; x)$ and ${}_1F_0(\cdot, \cdot; x)$ respectively are hyper-geometric and generalized hyper-geometric series.

The empirical coherence $\hat{\rho}$, its module d and phase φ are the ML estimators of degree of coherence ρ , coherence D and β , respectively. The distribution of the single-look phase difference is shown in Figure 5.14. For zero coherence, the phase difference is uniformly distributed. In both interferometry and radar polarimetry, any estimation of β only based on the phase difference has to be made on a high number of looks (or pixels) unless the coherence D is very close to 1. Aside from variance being generally high, bias can also arise from a calculation problem. Indeed, given the modulo 2π definition of phase, $E(\varphi) \neq \beta$, unless β is centered on the definition interval of φ , i.e., $\beta = 0$ with $\varphi \in [-\pi, \pi]$. In the case shown in Figure 5.14, for example, and $\varphi \in [-\pi, \pi]$, the mean phase differences would be 0° while $\beta = 180^\circ$. The distribution mode helps to avert this problem.

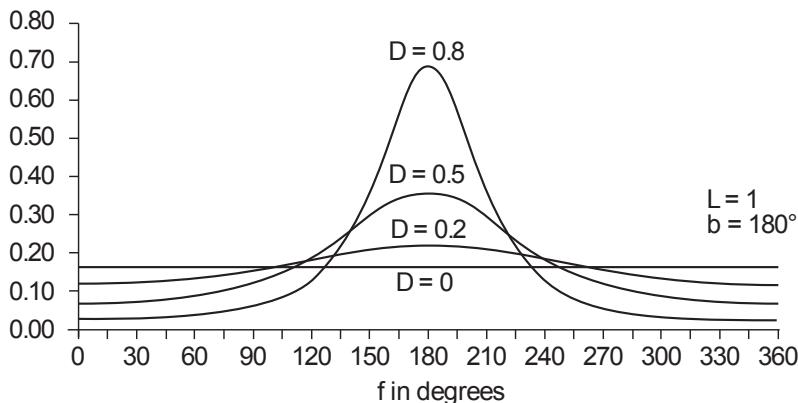


Figure 5.14. Single-look phase difference distribution for different values of coherence D

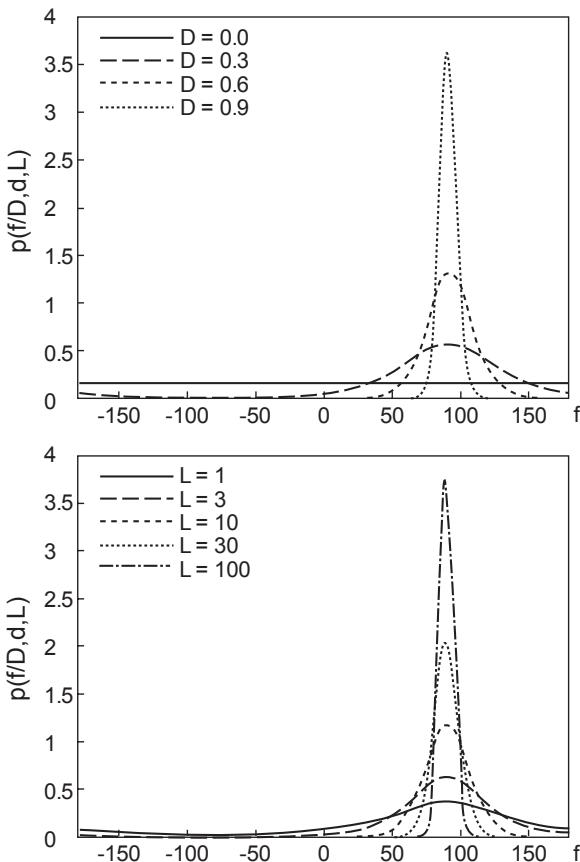


Figure 5.15. Multi-look phase difference distribution for a varying degree of coherence and $L = 10$ (top), and for a degree of coherence $D = 0.6$ and different values of L (bottom)

This shows we should not calculate the mean phase difference directly in interferometry, but rather via the Hermitian product argument, which is a multi-look spatial product in which single-look phase differences are weighted by amplitude A . Indeed, we have:

$$E(\hat{\rho}) = D(1-D^2)L \frac{\Gamma^2(L+\frac{1}{2})}{\Gamma(L)\Gamma(L+1)} {}_2F_1\left(L+\frac{1}{2}, L+\frac{1}{2}; L+1; D^2\right) e^{j\beta}.$$

The phase of the coherence degree or Hermitian product is therefore unbiased. Moreover, the estimation variance decreases as the number of looks (or pixels) L increases, as seen in both cases in Figure 5.15. The Cramer-Rao bound of the phase estimation variance is given by the following expression which only holds for high values of L and D :

$$\lambda_{CR} = \frac{1-D^2}{2LD^2}$$

However, we again stumble on the same difficulty described above. We therefore should not try to average phases of the coherence degrees or Hermitian products because the problem remains, but again go through an intermediate stage and calculate the mean of the available Hermitian product samples. In radar polarimetry, we will use, for example, a weighted mean of empirical covariance matrices Σ_z .

Unlike the Hermitian product phase, a multi-look product provides a biased estimation of coherence as:

$$E(d) = \left[\Gamma(3/2)\Gamma(L)/\Gamma(L+1/2) \right] (1-D^2)^L {}_3F_2 \left(3/2, L, L; L+1/2, 1; D^2 \right).$$

$E(d)$ differs from D , except where we have $D = 1$, or an infinite number of looks.

This bias is particularly high for low values of coherence D and low numbers of looks L , as can be seen in Figure 5.16. By contrast, the bias is almost negligible where $L \geq 20$ [BRU 96, TOU 99].

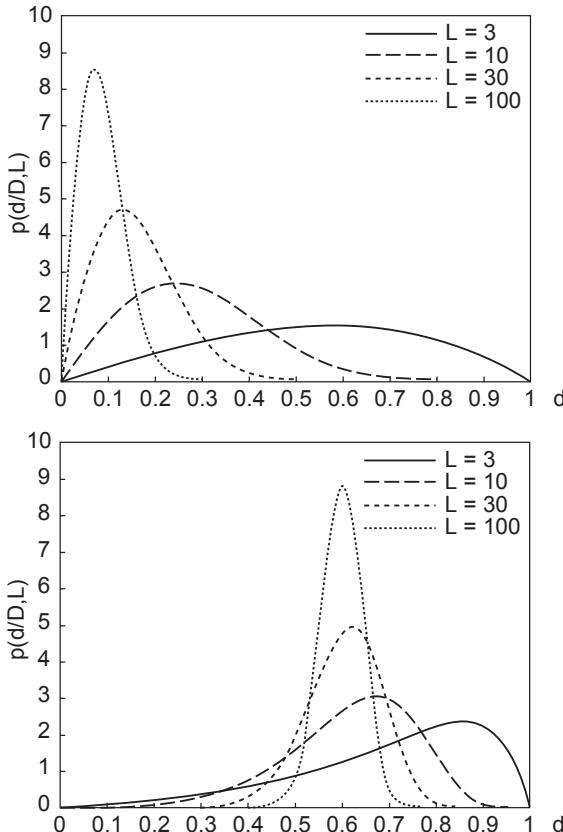


Figure 5.16. Empirical coherence distribution d for different values of the number of looks L and true coherence $D = 0$ (top) and for different values of looks L and true coherence $D = 0.6$ (bottom)

The bias and variance of the ML estimator decrease with the increase in D and are not zero except for $D = 1$. In any case, the estimation variance cannot be lower than the Cramer-Rao bound given by:

$$\frac{(1-D^2)^2}{2L}.$$

Different unbiased estimation methods of coherence D are reviewed in [TOU 99].

Finally, the joint distribution of two intensities makes it possible to deduce the distribution of ratio r between the operating channels according to L , D and the reflectivity contrast defined as $c = \frac{R_1}{R_2}$. The intensity ratio is the ML contrast estimator;

– joint distribution of correlated intensities:

$$p(I_1, I_2 | R_1, R_2, D, L) = \begin{cases} \frac{L^{L+1}}{R_1 R_2 \Gamma(L)(1-D^2)} \left[\frac{I_1}{R_1} \frac{I_2}{R_2} \right]^{(L-1)/2} \\ \cdot \exp \left[-\frac{L}{1-D^2} \left(\frac{I_1}{R_1} + \frac{I_2}{R_2} \right) \right] I_{L-1} \left[\frac{2LD}{1-D^2} \sqrt{\frac{I_1}{R_1} \frac{I_2}{R_2}} \right]; \end{cases}$$

– distribution of ratio $r = \frac{I_1}{I_2}$:

$$p_r(r | D, L, c) = c^L \frac{\Gamma(2L)}{\Gamma^2(L)} (1-D^2)^L \frac{(r+c)r^{L-1}}{\left[(r+c)^2 - 4D^2 cr \right]^{L+0.5}}.$$

Figure 5.17 shows the distribution of ratio r_0 for a 0 dB ($c = 1$) contrast and different values of D and L . If $c \neq 1$, the ratio is simply expressed as $r = cr_0$ and its properties are easily inferred from those of ratio r_0 . The r distribution for zero coherence D and different values of contrast c is given in Figure 8.4.

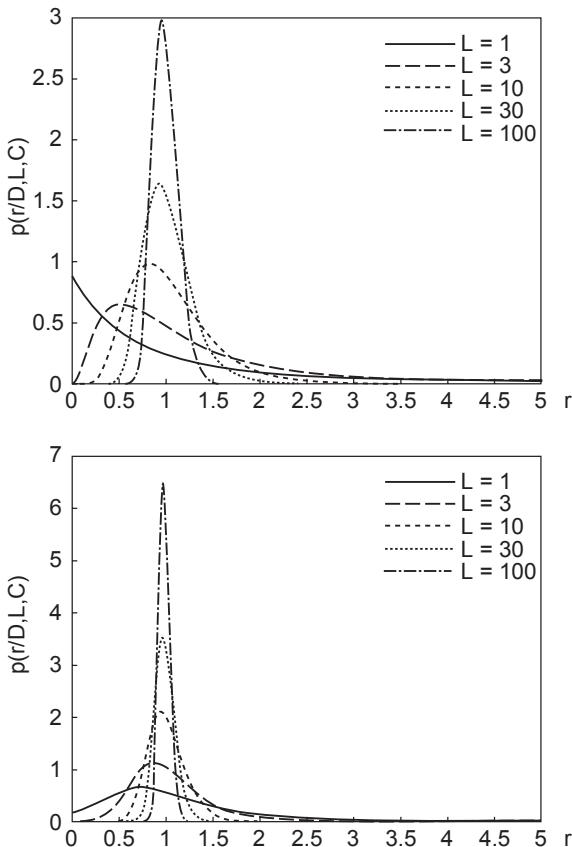


Figure 5.17. Distribution of intensity ratio r_0 for contrast $c = 1$ and coherence $D = 0.3$ (top), and for contrast $c = 1$ and coherence $D = 0.6$ (bottom)

r_0 moments are:

$$E(r_0^m) = (1 - D^2) \frac{\Gamma(L+m)\Gamma(L-m)}{\Gamma^2(L)} {}_2F_1(L+m, L-m; L; D^2)$$

The expectation of ratio $E(r) = c \frac{L}{L-1} \left(1 - \frac{D^2}{L}\right)$ shows that this is a biased estimation of contrast c by a higher value. The overestimation is maximal for zero coherence D , where the maximum bias given by $E(r) = c \frac{L}{L-1}$ is equivalent to 3 dB

for $L = 2$ looks, and 1.76 dB for $L = 3$ looks. The expectation is not defined for $L = 1$ look. The variance of unbiased estimator $\hat{c} = r \frac{L-1}{L-D^2}$ is:

$$\text{var}(\hat{c}) = c^2 \left[\frac{(L-1)(L-2)}{L(L-2)} \left(1 - \frac{D^2}{L} \right)^{-2} \left\{ 1 - \frac{2D^2}{L} \left(2 - \frac{3D^2}{L+1} \right) \right\} - 1 \right].$$

It reaches a maximum for $D = 0$ and decreases if either L or D go up. However, it is not zero unless $D = 1$. When coherence drops, we have to use a very large number of looks (or average over a large number of pixels), if we are to obtain a fair estimate of contrast, particularly in boundary detection (see Chapter 8), between remote dates in multi-temporal data or for polarization ratios between a co-polarized channel HH or VV and a cross-polarization channel HV or VH. The bias problem, however, is unimportant where ratio thresholding makes it possible to make a decision (e.g., in the case of a temporal change in radar reflectivity or in the presence of a radiometric boundary). It is very important in this case to be able to accurately calculate the estimation errors deriving from the choice of a particular threshold. Thus, the error likelihood, i.e., the likelihood that $r \notin [c(1-\eta), c(1+\eta)]$, may be calculated using the following relation:

$$P_E = \int_0^{c(1-\eta)} p_r(r|c, L, D) dr + \int_{c(1-\eta)}^{-\infty} p_r(r|c, L, D) dr.$$

This is maximum where correlation $\rho_l = D^2$ of two variables I_1 and I_2 is zero and does not depend on contrast c . Its value is then:

$$P_E = 1 + \frac{1}{L} \frac{\Gamma(2L)}{\Gamma^2(L)} \left[(1-\eta)^L {}_2F_1(2L, L; L+1; -(1-\eta)) \right. \\ \left. - (1+\eta)^L {}_2F_1(2L, L; L+1; -(1+\eta)) \right]$$

The first term in brackets corresponds to the first integral while all the others correspond to the second integral. They may be used, by thresholding the ratio with thresholds equal to $c(1 \pm \eta)$, to calculate likelihoods of detection and false alarms. Figure 5.18 indicates that working on a single pixel for routine numbers of looks ($L < 10$) leads to doubtful results. As with the case of reflectivity estimates, we have

to use the mean of a large number of pixels to estimate the contrast and thus increase the effective number of L looks. On the contrary, when correlation ρ_l tends toward 1, the ratio becomes a reliable contrast estimator even for low numbers of looks.

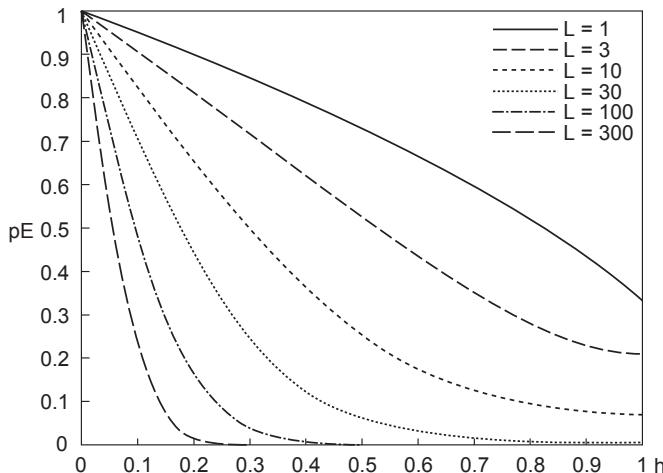


Figure 5.18. Maximum probability of error in estimating contrast c via the intensity ratio obtained for the worst case scenario in which correlation is zero

5.4. Non-Gaussian speckle models

Non-Gaussian speckle is found when the Goodman model conditions are no longer met at every pixel scale, or if they are still met, when the surface is no longer stationary. In the former case, this is essentially due to a small number of scatterers (or the predominance of a few scatterers) and low roughness, as may be seen, respectively, on images of urban sites captured by sensors with very high spatial resolution and in the nearly specular reflection produced, regardless of the spatial resolution, by smooth surfaces normal to the line of sight. When the surfaces, while still satisfying the Goodman criteria, are no longer broadly stationary, the observed phenomenon is often referred to as “texture”. We will also call such surfaces “heterogenous”. Within the framework of stationary heterogenous surfaces of second order, the scalar product model may be used to model the properties of such surfaces in statistical terms. These models may also be used to reduce speckle, describe target detection clutter and for classification purposes.

Finally, we can also model the data without any consideration of the multiplicative model. This makes it possible to resolve detection, segmentation and classification problems based on the initial data, without having to resort to

reflectivity R as an intermediary. In this case, the radar people would rather use the general term of clutter to characterize the response of the surfaces.

5.4.1. Scalar product model and normalized texture

5.4.1.1. Normalized texture variable

Consider the image of a surface in which the reflectivity of every pixel is a random variable $R(u)$ due to variation in the number of scatterers N_d and/or because the scatterers are not exactly the same. Let us assume that a stationary surface of the second order is characterized by the statistical properties of $R(u)$, especially its marginal distribution $p_R(R)$. The one-dimensional multiplicative speckle model introduced in section 5.2.2.3 may be rewritten as a 3-variable product in intensity, for example, as:

$$I(u) = R(u)S_I(u) = E(R)\frac{R(u)}{E(R)}S_I(u) = E(R)\mu(u)S_I(u)$$

Random variable $\mu(u)$ is the normalized *texture* of the surface, while $E(R)$ is its radiometry or mean radar reflectivity that is proportional to backscattering coefficient σ^0 used by physicists. The normalized texture consequently has unit mean and may be characterized by its marginal distribution $p_\mu(\mu)$. Distribution of μ is simply the distribution of reflectivity R with mean 1. More generally, these two variables may be described by second order statistics, or better still by the joint distribution of N spatial realizations of R .

Moreover, the multiplicative model assumes that speckle is independent of the scene, i.e., $E(R^m S_I^m) = E(R^m)E(S_I^m)$. This is of course an approximation, since S_I is the spatial organization of scatterers inside each resolution cell. However, as this microscopic property cannot be measured, at least given the spatial resolution of our sensor, it can be separated from R , which is the only measurable characteristic of the scene.

Note finally that this model, in which texture is represented by a multiplicative positive scalar variable, can easily be extended not only to detected amplitude A data, but also, under certain conditions, to complex multi-dimensional data, particularly the polarimetric backscattering vector \vec{S} (see section 5.5.2). The product model intensity equivalent is given by the expression of covariance \mathbf{C}_z and observed covariance Σ_z matrices in their respective forms [YUE 89, LOP 97b]:

$$\mathbf{C}_z(u) = \mu(u)\mathbf{C}_{zh} \quad \Sigma_z(u) = \mu(u)\Sigma_{zh}(u)$$

where \mathbf{C}_{zh} , i.e. the equivalent of $E(R)$ in the one-dimensional model, is the covariance matrix of an ideal, perfectly homogenous region, where the observed covariance matrices would be $\Sigma_{zh}(u)$. This model naturally cannot be applied to a set of N neighboring pixels. By contrast, it can be applied to interferometric images on surfaces where neither the physical nor the geometric parameters change significantly leading to texture changes from one date to the next.

5.4.1.2. Statistical properties of heterogenous area data

Let us consider the one-dimensional case. The marginal distribution of intensity I is easily obtained by integrating conditional distribution $p_I(I|R)$ over distribution $P_R(R)$:

$$p(I) = \int_0^{+\infty} p_I(I|R)p_R(R)dR$$

The distribution of I can also be obtained by considering the distribution of product $P = I = xy$, where $x = R$ and $y = S_I$ are random independent variables. This distribution is given by relation $p(P) = \int p_z(z)p_y(\frac{P}{z})\frac{1}{z}dz$, which in our case is:

$$p(I) = \int_0^{+\infty} p_R(z)p_{S_I}\left(\frac{I}{z}|z\right)\frac{1}{z}dz.$$

It is important that we identify this relation as the Mellin convolution for distributions of each variable [NIC 97, NIC 98]. Using the Mellin transform tables, we will then be able to deduce intensity distributions in various cases, some of which are given under the Pearson system for reflectivity R further below.

Based on the multiplicative model, we can also easily demonstrate that:

$$\begin{aligned} E(I) &= E(S_I)E(\mu)E(R) = E(R) \\ \text{var}(I) &= E(R)^2 \left[E(S_I^2)E(\mu^2) - 1 \right] \\ &= E(R^2) \left[(E(\mu^2) - 1)(E(S_I^2) - 1) + (E(S_I^2) - 1) + (E(\mu^2) - 1) \right] \end{aligned}$$

which will allow us to find the intensity coefficient of variation γ_I according to variables μ and S_I [ULA 86b]:

$$\gamma_I^2 = \gamma_{S_I}^2 + \gamma_\mu^2 + \gamma_{S_I}^2 \gamma_\mu^2.$$

This relation is widely used in spatial estimators of reflectivity R .

γ_I is equivalent to the standard deviation of luminance that is used in incoherent imaging to quantify surface heterogeneity in classifications. This can be used as an indicator in texture segmentations. Its lowest value is that of speckle alone $\gamma_I = \frac{1}{\sqrt{L}}$.

Observed on homogenous surfaces, it consequently grows as surface reflectivity becomes increasingly heterogeneous and can easily be generalized to amplitude data.

In practice, the coefficient of variation $\widehat{\gamma_I}$ is estimated from the ratio of the empirical standard deviation to the empirical mean. It also becomes a random variable as a result of speckle. Its distribution has not been established, but its first moments have been calculated for a homogenous surface [CAV 98]. When it is calculated on N independent samples, we have:

$$E(\widehat{\gamma_I^2}) = \frac{1}{L+\frac{1}{N}} \quad \text{var}(\widehat{\gamma_I^2}) = \frac{2L(L+1)}{(N-1)\left(L+\frac{1}{N}\right)^2\left(L+\frac{2}{N}\right)\left(L+\frac{3}{N}\right)}$$

The above relations help establish the confidence rate we can assign to a thresholding of the square coefficient of observed variation in order to determine the non-homogeneity of an area of N pixels. If N is high enough, we can consider that $\widehat{\gamma_{S_I}}$ is distributed according to a normal distribution allowing us to calculate the error rate using error function $\text{erf}(x)$ [LOP 93b].

Finally, based on the multiplicative model, we can also show that the auto-covariance function of intensity $C_I(\Delta u)$ is equal to the product of the speckle and texture auto-covariance functions: $C_I(\Delta u) = C_\mu(\Delta u) C_{S_I}(\Delta u)$. A relation similar to the one connecting coefficient of variations is then established between texture spatial correlation coefficient $\rho_\mu(\Delta u)$ and the correlation coefficients of intensity $\rho_I(\Delta u)$ and speckle $\rho_{S_I}(\Delta u)$:

$$\rho_I(\Delta u) \gamma_I^2 = \rho_{S_I}(\Delta u) \gamma_{S_I}^2 + \rho_\mu(\Delta u) \gamma_\mu^2 + \rho_\mu(\Delta u) \rho_{S_I}(\Delta u) \gamma_{S_I}^2 \gamma_\mu^2$$

However, this relation does not hold unless the sensor resolution is lower than texture correlation lengths [BLA 93].

When calculated on N independent samples, spatial correlation $\hat{\rho}$ provided by the conventional estimator has an estimation variance given by [RIG 93]:

$$var(\hat{\rho}) = \frac{1}{N} (1 - \rho)^2 (1 + 6\rho + 2\rho^2)$$

Consequently, a very large number of samples will be necessary to accurately estimate weak correlations as in the case of coherence D .

5.4.1.3. The Pearson system and K, U, B and W distributions

K distribution

Numerous pieces of research focused on the $p(R)$ distribution in order to describe the distribution of radar reflectivity in a natural scene as accurately as possible. Among the most interesting was Gamma distribution, which emerged as one of the most consistent distributions in the adjustment tests and one of the most justified in theory, as it relies on biophysical considerations where vegetation is concerned. The variation of scatterer number N_d can be modeled with the help of a process called “birth-death-migration”, which accounts for the natural evolution of vegetation. The number of scatterers in this case obeys a negative binomial distribution [JAO 84], where properties are reviewed in polarimetric speckle in section 5.5.2. The Gamma distribution, which is considered a distribution analog to the negative binomial distribution in a continuum, is therefore fully justified. The Gamma distribution characterized by scale parameter $E(R)$ and shape parameter ν is given by:

$$p_R(R) = \frac{1}{\Gamma(\nu)} \left[\frac{\nu}{E(R)} \right]^\nu \exp\left(\frac{-\nu R}{E(R)} \right) R^{\nu-1}$$

where the heterogeneity coefficient is:

$$\nu = 1/\gamma_R^2 = 1/\gamma_\mu^2.$$

A homogenous zone corresponds to the specific case where ν tends toward infinity. It leads to the following K distribution [JAK 76]:

$$P(I) = \frac{2\nu L}{\Gamma(\nu)\Gamma(L)E(R)} \left(\frac{\nu LI}{E(R)} \right)^{(\nu+L-2)/2} K_{\nu-L} \left(2\sqrt{\frac{\nu LI}{E(R)}} \right) \quad [5.14]$$

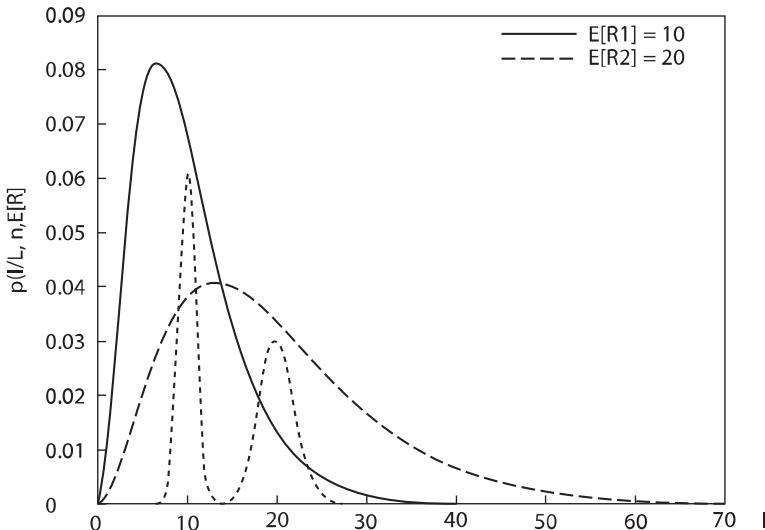


Figure 5.19. K-3 look distribution for two surfaces with 3 dB contrast on the mean reflectivities and a heterogeneity parameter of 100

Figure 5.19 illustrates the K-3 look distribution and underlying Gamma distribution $P_R(R)$. The assignment function of K-1 look distribution can be used to calculate the likelihood of false alarms caused by one pixel thresholding during target detection in a clutter. This function is given by:

$$Prob(I > T) = \frac{2}{\Gamma(\nu)} \sqrt{\frac{\nu T}{E(R)}} K_\nu \left(2\sqrt{\frac{\nu T}{E(R)}} \right)$$

The moments of order m are given by the following relation:

$$m_m = E(I^m) = R^m \frac{\Gamma(L+m)\Gamma(\nu+m)}{L^m\Gamma(L)\nu^m\Gamma(\nu)},$$

which allows the estimation of scale and form parameters, $E(R)$ and ν , respectively, based on empirical moments \widehat{m}_m calculated on N independent samples. Since the precise ML estimate cannot be obtained, we can simply use the intensity moments of first and second orders, i.e.:

$$\widehat{E(R)} = \bar{I} \quad \hat{\nu} = \frac{L+1}{L\left(\frac{\widehat{m}_2}{\widehat{m}_1} - 1\right) - 1}$$

where $\widehat{m}_1 = \bar{I}$ is the arithmetic mean of the intensities. To increase precision, albeit at the expense of numerical resolution, it is advisable to use non-integer moments [JOU 93, ISK 99]. On the other hand, if a high number of looks is available, the ML estimators of mean $E(R)$ and form parameter are better approximated by the following solution [OLI 93]:

$$\widehat{E(R)} = \widehat{m}_1 \left[1 + \frac{1}{L} \left(1 + \hat{\nu} - \hat{\nu} \frac{\widehat{m}_2}{\widehat{m}_1} \right) \right]$$

$$\ln \hat{\nu} - \Psi(\hat{\nu}) + \frac{\hat{\nu}}{L} \left[\frac{\widehat{m}_2}{\widehat{m}_1} - 1 \right] = \ln \bar{I}_1 - \ln \bar{I} + \ln \frac{1}{2L}$$

where $\Psi(x)$ is the digamma function. Variable $U = \ln \bar{I} - \ln \bar{I}_1$ in the right-hand member is the log of the arithmetic and geometric mean intensity ratio. It should be zero for a speckle-free homogenous surface ($I = \text{constant}$) and grow with heterogeneity. Therefore, like the empirical coefficient of variation, it turns out to be an indicator of heterogeneity particularly well suited for SAR imaging [OLI 00].

Of all distribution laws, the K distribution is currently the most widely used to characterize heterogenous surfaces. The two parameters related to the distribution of

R are supplemented by those of the spatial correlation coefficient $\rho_\mu(\Delta u)$ that help characterize every surface more fully, which is generally an exponential or Gaussian decrease modulated by a cosine term in radial direction to account for the negative correlation arising as a result of shadow effects or periodicity [OLI 86, POS 93].

Correlation coefficient $\rho_\mu(\Delta u)$ can be estimated from the relationship between the intensity correlation coefficients given at the end of section 5.4.1.2 where $\rho_I(\Delta u)$ is provided by the traditional estimator. To achieve yet higher accuracy, we can estimate spatial correlations from amplitude A data or logarithmic data [LOM 94].

The Pearson system and U, B, W distributions

Besides the Gamma distribution, other distributions, including Beta distribution of the first kind, have been advanced within the model's framework to find the distribution best suited for a given surface [LOP 90a]. As it happens, both the Gamma and Beta distributions are part of the Pearson system, which is a family of distributions that allow a broader, suppler parameterization, capable of providing a good description of marine clutter [DEL 97a]. The set of distributions of interest to us are a subsystem of the Pearson system, where probability density $p_R(R)$ is the solution of the following differential equation:

$$\frac{dp_R(R)}{dR} = -\frac{a + R}{R(c_1 + c_2 R)}$$

where a , c_1 , and c_2 are real parameters. Depending on their values, we will find the following solutions:

- if $c_2 = 0$, we obtain a 3rd type (Gamma distribution) Pearson function and K distribution is the solution (equation [5.14]);
- if $c_1 = 0$ and $c_2 > 0$, we are in the situation of a 5th type (“inverse Gamma” function) Pearson function and the solution is written (type B as Beta distribution of the second kind):

$$P_I(I) = \frac{\Gamma(\alpha + L)}{\Gamma(\alpha)\Gamma(L)} \left(\frac{LI}{\xi} \right)^L \frac{1}{I} \frac{1}{\left(1 + \frac{LI}{\xi} \right)^{L+\alpha}}$$

- if $c_2 < 0$ and $c \neq 0$, we have a 1st type Pearson distribution (Beta distribution of the first kind), and the solution (W type) is:

$$P_I(I) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(L)} \left(\frac{LI}{\xi} \right)^{\frac{p+L-1}{2}} \frac{1}{I} e^{-\frac{LI}{2\xi}} W_{\frac{1-2q-p+L}{2}, \frac{p-L}{2}} \left(\frac{LI}{\xi} \right)$$

where W is the Whittaker function;

– if $c_2 > 0$ and $c \neq 0$, this is the case of a 6th type Pearson function (Beta function of the second kind). The solution (U or Psi type) is:

$$P_I(I) = \frac{\Gamma(p+q)\Gamma(q+L)}{\Gamma(L)\Gamma(p)\Gamma(q)} \frac{1}{I} \left(\frac{LI}{\xi} \right)^p \Psi \left(p+q, 1-L+p; \frac{LI}{\xi} \right)$$

where $\Psi(a, b; x) = \frac{1}{\Gamma(a)} U(a, b; x)$ is a confluent hyper-geometric function, ζ is the scale parameter and the others are form parameters.

The amplitude distributions are easily deduced from the intensity distribution laws, according to $p_A(A) = 2Ap_I(A^2)$. Several methods to estimate the parameters of these laws are given in [DEL 93, MAR 96, DEL 97b].

The U and W distributions, corresponding to 3-parameter rather than 2-parameter distributions, as do the other distributions, help describe a broader variety of texture distribution shapes, but lack the theoretical grounding of the K distribution. A generalized inverse Gaussian distribution has also been advanced recently. The main advantage of this 3-parameter distribution lies in the fact that it includes the Gamma and inverse Gamma distributions as particular cases and provides a better description of extremely heterogenous clutters [FRE 97]. Finally, an extension of the multiplicative model to the multi-dimensional case presented in the polarimetric section can be approached within the broader framework of spherically invariant random processes and vectors [CON 87, RAN 93].

5.4.2. Non-Gaussian clutter

Within the framework of the multiplicative model, clutter is regarded as essentially Gaussian: the conditional distribution of each complex pixel value $z = i + jq$, given reflectivity R , is Gaussian, even if all the distributions of the derived variables, particularly those mentioned in the previous section, are no longer so. However, we do not always need to separate Gaussian speckle from non-Gaussian reflectivity. Finally, speckle may in some cases be non-Gaussian (not fully developed), e.g., when there are few scatterers. Many empirical distribution laws

have therefore been advanced to globally model radar data clutter outside the Gaussian process.

Those that have been successfully applied to radar images include the Pearson system distributions directly assigned to data for classification or segmentation purposes [MUR 87, MAF 91, DEL 97b], generalized Gamma distributions, lognormal distributions, and those of Weibull and Nagakami-Rice. In the following, we will briefly present the latter three that have not been discussed above as well as the partially developed speckle.

5.4.2.1. The lognormal distribution

The lognormal distribution applies to an item of data x (I or A) when its logarithm is distributed according to a Gaussian distribution. It is written:

$$P_x(x) = \frac{1}{x\sqrt{2\pi\sigma_L^2}} \exp\left[-\frac{(\log x - m_L)^2}{2\sigma_L^2}\right]$$

where m_L and σ_L^2 are the expectation and variance of $\log x$. Its moments are given by $E(x^m) = E^m(\log x) \exp\left[\frac{m(m-1)}{2}\sigma_L^2\right]$ where $E(x) = \exp\left(m_L + \frac{1}{2}\sigma_L^2\right)$. This distribution is best suited for amplitude data and urban surfaces, and generally for very heterogeneous surfaces with many powerful scatterers. The ML parameter estimators are given in [SZA 77] and [DEL 97a].

5.4.2.2. Weibull distribution

The Weibull distribution is another two-parameter distribution given by $P_x(x) = \left(\frac{x}{\xi}\right)^\eta \frac{\eta}{\xi} \exp\left(-\left(\frac{x}{\xi}\right)^\eta\right)$ where ξ is the scale parameter and η the shape parameter. Its moments are given by relation:

$$E(x^m) = \xi^m \frac{\Gamma\left(1 + \frac{m}{\eta}\right)}{\Gamma^m\left(1 + \frac{1}{\eta}\right)}$$

and its assignment function is $P(x > T) = 1 - \exp\left[-\left(\frac{T}{\xi}\right)^\eta\right]$. When $\eta = 1$, we are back to the exponential distribution (single-look intensity distribution on an

homogenous surface) and when $\eta = 2$, what we have is the Rayleigh distribution (single-look amplitude distribution on an homogenous surface), which may explain its use for weak heterogenities. It is thus not very well suited for an overall modeling of multi-look data and strong heterogenities. Estimators of these parameters are given in [MEN 63] and [OLI93].

5.4.2.3. Nagakami-Rice distribution

When we have a single very powerful scatterer with RCS $\sigma^o = A_o^2$ immersed in homogenous clutter with reflectivity R , we can determine the distribution of the backscattered intensity $x = I$ for single-look images by relation:

$$p_I(I) = \frac{1}{R} e^{-\frac{I+A_o^2}{R}} I_0\left(2\frac{\sqrt{IA_o^2}}{R}\right)$$

This distribution can be used to detect a fixed target in a Gaussian clutter. The assignment function, which here enables us to calculate the detection probability, is given by [DAN 86]. The ML parameter estimators are given in [SIJ 98].

5.4.2.4. The partially developed speckle

When the number of scatterers N_d is low (lower than 10), but the phase differences are still uniformly spread and independent of amplitudes in order to represent rough surfaces, the speckle distribution law becomes hard to model. Nevertheless, in a few particular cases where $N_d = 2.3$ and basic amplitude a_n remains constant, the distribution law has been established within the framework of the two-dimensional random walk and from approximations where $N_d > 3$ [DAB 95]. In other cases, the intensity moments can be linked to the moments of random processes N_d and a_n . If, for example, we neglect the effect of the sensor's impulse response, the first two moments are given by $E(I^2) = E(N_d)E(a^2)$ and:

$$E(I^2) = E(I)^2 \left[\frac{2E(N_d(N_d - 1))}{E(N_d)^2} + \frac{E(a^4)}{E(N_d)E(a^2)^2} \right],$$

whereas for Gaussian speckle, $E(I^2) = 2E(I)^2$ and for K-distributed speckle, $E(I^2) = 2E(I)^2 \left[1 + \frac{1}{\nu}\right]$. These relations between moments help us obtain more accurate information on either of the two processes in the case in which the properties of one of them are known beforehand.

However, partially developed speckle corresponds to sensors with spatial resolution that is close to wavelength, and this is where the concept of point scatterers should give way to more physical approaches to backscattering. SAR images in the VHF field and at frequencies lower than 90 MHz recently made available have revived interest in these issues [JAK 84, GOO 86] and [BEC 87].

5.5. Polarimetric radar speckle

5.5.1. The Gaussian model

In order to derive the Gaussian model, the central limit theorem is applied to estimate the effect resulting from the vector sum of basic effects of each basic scatterer inside the resolution cell, assuming they are numerous. To make this application possible, several conditions have to be met: we particularly need to have a fully developed speckle on a homogenous (untextured) stationary scene. The components of $\vec{\mathbf{S}}^1$ are then complex Gaussian random variables with 0 mean:

$$E(\operatorname{Re}(S_{ij})) = E(\operatorname{Im}(S_{ij})) = 0$$

and: $E(\operatorname{Re}(S_{ij}^2)) = E(\operatorname{Im}(S_{ij}^2))$ (circular speckle)

$$E(\operatorname{Re}(S_{ij})\operatorname{Im}(S_{ij})) = 0 \quad \text{(decorrelation of real and imaginary parts)}$$

$\vec{\mathbf{S}}$, with the dimension d , follows a centered Gaussian distribution:

$$P(\vec{\mathbf{S}}) = \frac{1}{\pi^d |\mathbf{C}|} \exp(-\vec{\mathbf{S}}^{*t} \mathbf{C}^{-1} \vec{\mathbf{S}}) \quad [5.15]$$

where $\vec{\mathbf{S}}^{*t}$ is the conjugate transpose of $\vec{\mathbf{S}}$, $|\mathbf{C}|$ is the determinant of \mathbf{C} (covariance matrix of $\vec{\mathbf{S}}$; see equation [1.19]) and d is the dimension of $\vec{\mathbf{S}}$, leading to 3 or 4.

1 We have kept here for the polarimetric quantities the same notations as in Chapter 1: \mathbf{S} is vector \mathbf{Z} found in the previous sections of this chapter.

Therefore, the distribution is fully characterized by covariance matrix \mathbf{C}^2 , from which all higher order moments are deduced. In particular, odd-order moments are zero and the non-zero moments of order 4 are given by:

$$\begin{aligned} E(S_i S_j S_k^* S_l^*) &= E(S_i S_j) E(S_k^* S_l^*) + E(S_i S_k^*) E(S_j S_l^*) + E(S_i S_l^*) E(S_j S_k^*) \\ &= C_{ik} C_{jl} + C_{il} \cdot C_{jk} \end{aligned} \quad [5.16]$$

because by assuming an independent backscattering of the various backscatterers inside the resolution cell, we have $E(S_i S_j) = 0$ (uniformly distributed phase of the backscattering elements).

According to equation [5.16], we obtain in particular: $E(S_i S_j S_i^* S_j^*) = C_{ii} C_{jj} + C_{ij} C_{ji} = E(I_i) E(I_j) + |C_{ij}|^2$.

Finally, note that for L -look multi-look data, the observed covariance matrix Σ follows a complex Wishart distribution law (equation [5.8]) [GOO 76a]:

$$p(\Sigma | \mathbf{C}) = \frac{L^d |\Sigma|^{L-d} \exp\{-L \cdot \text{Tr}(\mathbf{C}^{-1} \Sigma)\}}{\pi^{d(d-1)/2} \Gamma(L) \cdots \Gamma(L-d+1) |\mathbf{C}|^L}. \quad [5.17]$$

5.5.2. The product model

The Gaussian model is by far the most widely used. Theoretically, however, it does not hold unless the number of scatterers inside the resolution cell is not only large, but also constant. If the number varies, texture develops. The general case is then rather that of a product model of two independent processes:

$$\vec{S} = \sqrt{\mu} \vec{G} \quad [5.18]$$

where \vec{G} is a multi-variate Gaussian random variable (with the same dimension as \vec{S}), and $\sqrt{\mu}$ is a scalar positive independent random variable representing texture. The target is then modeled by a high number of identical backscatterers, where the

² \mathbf{C}_z in previous sections.

exact number N_d varies from one cell to another. Therefore, for example, if N_d follows a binomial distribution [JAK 87], μ follows a Gamma distribution and we have a **K** distribution, though in multi-dimensional space this time, for $\bar{\mathbf{S}}$ [YUE 90]. Of course, $\mu = \text{constant}$ ($N = \text{constant}$) corresponds to the Gaussian case.

Moments of $\sqrt{\mu}$ are given as functions of N moments, which is the number of backscattering elements [OLI 98]:

$$E(u^k) = E(N_d^k)/E(N_d)^k \quad [5.19]$$

From this we especially deduce that the terms of the covariance matrix associated with $\tilde{\mathbf{G}}$ are equal to the terms of \mathbf{C} , which is the covariance matrix of $\bar{\mathbf{S}}$: $E(S_i S_j^*) = E(\mu) E(G_i G_j^*) = E(G_i G_j^*)$. For further details on the mathematical expansion and texture parameter estimates, see, for example, [TOU 95, LOP 97b].

Finally, an important comment about the product model that, by considering the different polarimetric channels, we get rid of texture (the μ parameter).

The validation of this product model was studied in [QUE 95]. The authors showed a Gaussian behavior of all types of cultures (in an agricultural site) at C-band; then they showed texture effects at longer wavelengths (L- and P-bands) and the need to resort to the K distribution. They singled out a number of problems, including the Gaussian behavior of the forest regardless of the band (although it is on this particular type of vegetation that the texture effect should first be found at longer wavelengths), the texture parameter variation with polarization (even though μ , which only depends on N_d , should be independent of polarization) at P-band in the grain crops, which is neither Gaussian nor K-distributed behavior of the town area (already noted by other authors).

5.5.3. Distributions of the phase difference and 2-amplitude ratio

The statistics of many parameters can be deduced based on the previous distributions. Here, we will only focus on the phase difference distributions between two channels and the ratio of two amplitudes.

We denote by ρ the correlation coefficient of channels i and j ($ij \in \{hh, hv, vh, vv\}$), with module $\rho = \frac{C_{ij}}{\sqrt{|C_{ii}| |C_{jj}|}}$, $|\rho| = D$ and phase β_ρ .

The probability density of phase difference θ_{ij} between channels i and j is then given by:

$$P(\theta_{ij} = \theta) = \frac{1}{2\pi} \times \frac{1 - |\rho|^2}{1 - y^2} \times \left\{ 1 + \frac{y}{\sqrt{1 - y^2}} \cos^{-1}(-y) \right\} \quad [5.20]$$

where $y = |\rho| \cos(\theta - \beta_\rho)$ and $-\pi \leq \theta \leq \pi$.

It is hence symmetric, single mode, centered at β_ρ and increasing with the decrease of $|\rho|$.

The phase difference between co-polar channels hh and vv is used to characterize some backscattering mechanisms [ULA 87, VAN 89] or [LET 92], especially for corn and forests. By contrast, this phase difference is assumed to have mean zero for data calibration (a good summary is given in [QUE 94]).

Figure 5.20a shows the image at L-band of phase difference θ_{hhvv} between co-polar channels hh and vv in the case of an agricultural site. The corresponding histogram is seen in Figure 5.20b. We note the high-level noise inherent in the image; the same image following speckle filtering can be seen in section 6.5. The histogram clearly illustrates the prevalence of zones with zero phase difference. It also shows very large phase deviations mostly arising from the different backscattering mechanisms that are at play in the H and V polarizations as a result of the geometries of some backscatterers such as stem.

The probability density of amplitude ratio $r_{ij} = |S_i/S_j|$ is given by:

$$P(r_{ij} = r) = 2c \cdot \frac{(1 - |\rho|^2)(c + r^2)r}{\left[(c + r^2)^2 - 4c|\rho|^2r^2 \right]^{3/2}} \quad [5.21]$$

where $c = E(|S_i|^2)/E(|S_j|^2)$.

These ratios are most widely used to remove surface geometric effects (e.g., relief, roughness). [OH 92] and [DUB 95], for instance, contemplate the use of the co-polar channel ratio to assess soil humidity.

5.5.4. Correlation of the co- and cross-polar channels

A simplifying assumption most widely used is the decorrelation of co- and cross-polar channels:

$$E(S_{hh}S_{hv}^*) = E(S_{hh}S_{vh}^*) = E(S_{hv}S_{vv}^*) = E(S_{vh}S_{vv}^*) = 0.$$

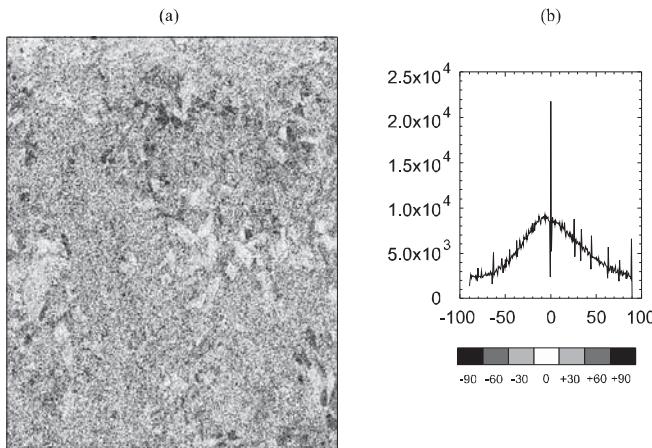


Figure 5.20. (a) Phase difference between co-polar channels hh and vv at L-band (Orgeval site, France); (b) phase difference histogram

Under this assumption, the number of terms characterizing the covariance matrix \mathbf{C} is cut to 4 real terms + 1 complex term, down from 4 real terms + 6 complex ones, or more commonly, under the reciprocity assumption, to 3 real terms (the 3 intensities: hh, hv and vv) + 1 complex term (the correlation between hh and vv), down from 3 real terms + 3 complex terms:

$$\mathbf{C} = \begin{pmatrix} E(|S_{hh}|^2) & 0 & E(S_{hh}S_{vv}^*) \\ 0 & E(|S_{hv}|^2) & 0 \\ E(S_{vv}S_{hh}^*) & 0 & E(|S_{vv}|^2) \end{pmatrix} \quad [5.22]$$

In fact, this hypothesis of a decorrelation of co- and cross-polar channels has only been demonstrated for the case of targets with cross-range symmetry, first according to a model expanded to the second order [BOR 87], then via symmetry based on invariance properties [CLO 95]. It is considered verified on most natural surfaces and used as a simplifying assumption in many applications. It provides the foundation for the calibration algorithms [VAN 89, KLE 92] used at JPL, for which it is applied over regions wide enough to be verified “on average”.

In the case of a Gaussian distribution, [GOO 76b] has linked (Siegeert relation) the complex correlation coefficient module ρ_{ij} of channels i and j ($i, j \in \{hh, hv, vh, vv\}$) to the intensity coefficient of these same channels (see section 5.3.2):

$$\rho_{ij}^{(I)} = \frac{E((I_i - E(I_i))(I_j - E(I_j)))}{\sqrt{E((I_i - E(I_i))^2)} \sqrt{E((I_j - E(I_j))^2)}} \quad \text{with: } I_i = |S_i|^2$$

$$\rho_{ij}^{(I)} = |\rho_{ij}|^2 \quad [5.23]$$

In the case of the product model and assuming $\tilde{\mathbf{G}}$ and μ are independent, the relation becomes:

$$\rho_{ij}^{(I)} = \frac{1}{\nu + 2} + \frac{\nu + 1}{\nu + 2} \times |\rho_{ij}|^2 \quad [5.24]$$

where ν is the Gamma distribution parameter of texture μ .

These relations are particularly useful when we try to estimate the correlation, especially of co-polar channels hh and vv, based on nothing but intensity images, as is the case of some filters described in section 6.5 (polarimetric data filtering).

In fact, this hypothesis of a decorrelation of co- and cross-polar channels has only been demonstrated for the case of targets with cross-range symmetry, first according to a model expanded to the second order [BOR 87], then via symmetry based on invariance properties [CLO 95]. It is considered verified on most natural surfaces and used as a simplifying assumption in many applications. It provides the foundation for the calibration algorithms [VAN 89, KLE 92] used at JPL, for which it is applied over regions wide enough to be verified “on average”.

In the case of a Gaussian distribution, [GOO 76b] has linked (Siegeert relation) the complex correlation coefficient module ρ_{ij} of channels i and j ($i, j \in \{hh, hv, vh, vv\}$) to the intensity coefficient of these same channels (see section 5.3.2):

$$\rho_{ij}^{(I)} = \frac{E((I_i - E(I_i))(I_j - E(I_j)))}{\sqrt{E((I_i - E(I_i))^2)} \sqrt{E((I_j - E(I_j))^2)}} \quad \text{with: } I_i = |S_i|^2$$

$$\rho_{ij}^{(I)} = |\rho_{ij}|^2 \quad [5.23]$$

In the case of the product model and assuming $\tilde{\mathbf{G}}$ and μ are independent, the relation becomes:

$$\rho_{ij}^{(I)} = \frac{1}{\nu + 2} + \frac{\nu + 1}{\nu + 2} \times |\rho_{ij}|^2 \quad [5.24]$$

where ν is the Gamma distribution parameter of texture μ .

These relations are particularly useful when we try to estimate the correlation, especially of co-polar channels hh and vv, based on nothing but intensity images, as is the case of some filters described in section 6.5 (polarimetric data filtering).

Chapter 6

Reflectivity Estimation and SAR Image Filtering

6.1. Introduction

Speckle, as seen in the previous chapter, greatly disrupts radar image readability. Due to this, numerous efforts were made to “filter” these images based on two main philosophies. One of them, historically the first to be adopted, aims to improve image appearance by limiting speckle. Standard filters (mean, median, etc.) were used to process SAR images in this approach with the clear goal of bringing radar imaging visually as close as possible to optical images, so that existing methods originally developed for the latter became applicable to the former. The focus then gradually extended from the mere visual aspect to the mathematical problem of estimating reflectivity, i.e. the physical parameter that is characteristic of these objects. “Filters” thus turned into “estimators” taking the specific nature of radar imaging into account. We have moved from noise removal to a different approach aimed at extracting information. The two are of course related, since the images of estimated reflectivity are ideally free of speckle. Further parameters, e.g., the covariance matrix in polarimetric imaging¹ and coherence in interferometry, can also be estimated. In a reflectivity estimation philosophy, the boundary between filtering and segmentation, or even classification, both of which will be discussed in the following chapters, becomes very thin. Indeed, we may find it useful to pair

Chapter written by Armand LOPÈS, Florence TUPIN and Sylvie LE HÉGARAT-MASCLE.

¹ A separate section will deal with this type of data, although they can be well considered as a specific case of complex multi-channel data.

different approaches and resort to, for example, segmentation or structure detection to improve our reflectivity estimates, as will be shown later in this chapter.

Radar image filtering is a complex problem, given the vast choice of schemes the user may be faced with. On the one hand, radar data come in a broad variety. They may be detected (amplitude, intensity), complex, polarimetric, single- or multi-look, multi-data, etc. On the other hand, when working with estimates, there is a wide variety of mathematical criteria that may come into play. A number of approaches complete with the different circumstances (types of data and types of estimators) to which they apply are described in this chapter.

We will first return to the standard image processing filters we have mentioned. Commercial software in the field frequently offers noise reduction filters designed for universal usage. Their specific action on radar images may therefore be worth knowing.

The simplest of them all is the mean filter that replaces the value of a pixel by the mean value of its neighbors. In radar imaging, such filters can be regarded as a spatial implementation of the multi-look technique, with the size of the pixel artificially kept unchanged. A square window is generally considered as neighborhood. As an effect of applying this filter, the homogenous areas become quite smooth, and speckle is reduced, but the edges are blurred. The larger the processing window, the more pronounced these effects are. In order to preserve sharp edges, we may use median filters that assign to the central window pixel the median value rather than the average one. This type of filter yields good results for the edges which stay crisp, but the homogenous areas become less smooth.

Other unspecific filters have also been applied such as the Lee σ filter based on a Gaussian hypothesis [LEE 83], mathematical morphology filters [SAF 89], Crimmins geometric filters [CRI 85, CRI 86], etc.

While these filters do attenuate the grainy noise, they also have some drawbacks. Most of them such as the median filter have a biasing effect on reflectivity, which is all the more marked as the number of looks is low (the mode is then indeed quite different from average; see Chapter 5). Filtering more often than not leads to some loss of information, especially in terms of textures which are destroyed in the process. Moreover, filters of this type address speckle as noise and therefore cannot exploit the information it contains.

6.2. Estimations of reflectivity R

This section deals with the estimators that make explicit use of the radar image formation process. Before we go on to describe at length some simple estimation filters for detected data (intensity and amplitude) and complex data, we will formally pose the problem of reflectivity estimation.

As seen in the previous description of speckle laws (Chapter 5), the physical parameter on which radar imaging relies is the backscattering ground factor. This parameter is proportional to the mean intensity of the scene, i.e. reflectivity R , which is the value we want to extract from radar images.

The problem facing us is therefore that of estimating R from the observations at hand, namely the values of intensity and amplitude or the complex value of the backscattered wave at every point of the image. This very standard problem may be considered within the framework of the Bayesian estimation, and we review its principles at this point, since they will frequently come up hereafter in this chapter.

6.2.1. Bayesian estimation recap

X being a hidden random variable, we are trying to determine its realization x (the solution sought). Y is a random variable, and we are observing a realization y . X and Y are related by their joint probability $P(X, Y)$, or conditional law $P(Y|X)$. We will estimate x by $\phi(y)$ from the observed y estimate. In our case, X represents reflectivity R and Y the observation (intensity, complex data, etc.) that is available to us. Also suppose that we know a cost function $L(x, \phi(y)) = L(x, x')$ that satisfies the relation:

$$L(x, x') \geq 0 \quad L(x, x') = 0 - x = x'$$

We define the optimal estimator with the optimal function ϕ as minimizing the cost expectation conditional to our knowledge of y , i.e.:

$$E(L(X, \phi(y))|Y = y) = \sum_{x \in \Omega} L(x, \phi(y)) P(x|y)$$

Function ϕ^{opt} thus minimizes the mean error arising from our substitution of $\phi(y)$ for x , conditional to the item of data y , and the “optimal” estimator is $\hat{x} = \phi^{\text{opt}}(y)$. We obtain different estimators depending on the cost functions considered. The

functions most widely used are the Kronecker and quadratic functions, for which we find the following results.

6.2.1.1. The Kronecker function: $L(x, x') = 1 - \delta_{x, x'}$

The optimal estimator is then:

$$x = \phi^{\text{opt}}(y) = \text{Argmax}_{\phi} [P(X = \phi(y)|y)]$$

It is the estimator that maximizes the *a posteriori* probability, or MAP. When we have no *a priori* knowledge of X , we can apply Bayes rule to derive the maximum likelihood (ML) estimator:

$$\text{Argmax}_{\phi} [P(Y = y|X = x)]$$

In order to calculate this estimator, we generally consider the logarithm of $P(X|Y)$ and find the x value that will cancel its derivative.

6.2.1.2. The quadratic cost function: $L(x, x') = (x - x')^2$

The optimal estimator in this case is:

$$\phi^{\text{opt}}(y) = E(X|y)$$

This estimator, which minimizes the mean square error, is known as the MMSE or MAP (maximum *a posteriori*) estimator. Because the expression to which it leads is often hard to resolve, suboptimal solutions need to be found. To find them, we will impose additional constraints on the shape of ϕ . The most common estimator requires that $\phi(\phi(Y)) = aY + b$ be linear, hence the acronym LMMSE. For an unbiased estimator, we will then have the following conditions:

$$E(\phi(Y)) = E(X) \quad [6.1]$$

$$\text{cov}(\phi(Y) - X, Y) = 0 \quad [6.2]$$

The above estimators (MAP, ML, MMSE, LMMSE) will be widely used later on. We will also single out the point case in which we estimate the value of R in a pixel where we only know the observation for that particular point, and the spatial or

contextual case in which we also use observations relative to the neighboring pixels. We could now note that the estimators we obtain bring into play various statistical quantities (mean, standard deviation, etc.) always empirically calculated over the neighborhood of the considered pixel. This aspect is discussed in section 6.6.

6.2.1.3. Notes on estimator behavior

We are providing here some of the criteria used to assess the quality of an estimator $\hat{X} = \phi(Y)$: bias, which measures deviation from the mean, estimator variance, which measures the dispersion of the estimators (minimal variance estimators are particularly interesting), and the standard deviation.

In the simple case where we estimate a deterministic parameter x , the estimator is considered to be unbiased if $E(\hat{X}) = x$. If X itself is a random variable (as it is here, since X represents reflectivity), we can distinguish two sorts of biases. First, there is a bias conditional on x (the observed value), which is $E_y(\hat{X}|X=x) - x$, where expectation E_y is calculated with conditional distribution $p(Y=y|X=x)$:

$$E_y(\hat{X}|X=x) - x = \int \phi(y) p(y|x) dy - x$$

In the pixel case, all estimators are conditionally biased except for the ML estimator that is equal to the datum itself (intensity I , as will be seen below).

The second type of bias is the “overall” or radiometric bias that integrates the probability density of X and is expressed by $E_x(E_y(\hat{X}|X)) - E_x(X)$, where E_x is the expectation with respect to $p_x(x)$:

$$\begin{aligned} E_x(E_y(\hat{X}|X)) - E_x(X) &= \iint \phi(y) p(y|x) p(x) dx dy - E_x(X) \\ &= E_y(\phi(y)) - E_x(X) \end{aligned}$$

Unlike the MAP (maximum *a posteriori*) estimator, the MMSE (minimum mean square error) estimators are unbiased in this case.

In order to carry out a complete study of the estimators, further criteria such as variance or standard deviation, which incorporates the bias effect, will need to be examined. These may be conditional or overall criteria.

We speak of *conditional* values when distribution $p(Y|X=x)$ occurs and *overall* values when distribution $p(X)$ is used. These different quantities (bias, variance, MSE, conditional and overall values) are used in section 6.7.1 in order to compare the various filters used in radar imaging.

6.2.2. Estimation of constant reflectivity R

We will first consider the simplest case, i.e. in which the reflectivity of a scene is supposed to be constant ($R = E(R)$, i.e. the scene is texture-free).

6.2.2.1. Multi-look filters

First consider the multi-look case. Suppose we have several looks of the same scene, i.e., several data for the same pixel, hence the same value of reflectivity. These different looks are generally obtained by dividing the across-track spectrum into L segments and creating an image for each segment (see Chapter 5). We will now examine the different types of data.

Complex data: for complex multi-look data in the form of a vector \mathbf{Z} , the distribution shape according to the speckle covariance matrix \mathbf{C}_S was given in Chapter 5. Its likelihood is written:

$$\log p(\mathbf{Z}|R) = -L(\log(\pi) = \log R) - \log|\mathbf{C}_S| - \frac{1}{R} \mathbf{Z}^{*t} \mathbf{C}_S^{-1} \mathbf{Z}$$

The ML estimator is therefore given by:

$$\hat{R} = \frac{1}{L} \mathbf{Z}^{*t} \mathbf{C}_S^{-1} \mathbf{Z}$$

which is an unbiased estimator. This filter, known as a whitening filter, was originally developed for polarimetric data (see section 6.5). The estimation of matrix \mathbf{C}_S will be addressed in section 6.6. Speckle reduction on the multi-look image is maximum in this case and the reflectivity estimate coefficient of variation is divided by \sqrt{L} .

Intensity data: let us first review the expression of the ML estimator in the case of single-look intensity. We obtain, according to [5.2]:

$$p(I|R) = \frac{1}{R} e^{-\frac{I}{R}}.$$

Hence:

$$\frac{d}{dR} \log(p(I|R)) = 0 - \frac{d}{dR} \left(-\log(R) - \frac{I}{R} \right) = 0$$

with a solution of $\hat{R} = I$. The ML estimator of reflectivity is simply the intensity at a certain pixel.

In the intensity multi-look case and assuming that looks I_1, \dots, I_L are independent, we obtain:

$$p(I_1, \text{etc., } I_L | R) = \prod_{i=1}^L p(I_i | R)$$

and using equation [5.2], the ML is written:

$$\log p(I_1, \text{etc., } I_L | R) = -\log R - \frac{1}{L} \sum_{i=1}^L I_i$$

The ML estimator for which the derivative is 0 then corresponds to the arithmetic average of intensities $\hat{R} = \frac{1}{L} \sum_{i=1}^L I_i$.

In the case of correlated intensity data, we do not have any joint expression of the probability density and consequently the ML estimator cannot be established. In practice, the L looks are averaged in intensity, but speckle is not reduced by a factor L , but rather by a factor $L' < L$ (see equation [5.6]).

Amplitude data: according to the same reasoning as above, we see that the ML estimator of reflectivity is again given by the arithmetic average of intensities $\hat{R} = \frac{1}{L} \sum_{i=1}^L A_i^2$. By contrast, the estimator $\left(\frac{1}{L} \sum_{i=1}^L A_i \right)^2$ is suboptimal [ZRN 75].

If complex data are available, it is more relevant to use a whitening filter rather than calculate the intensity and average over intensities. Indeed, if the pixels are correlated, the apparent number of looks L' will be lower than the real number L , while it will be equal to L for complex data. It is as though the whitening filter was decorrelating the pixels before adding incoherently, which leads to improved results for the constant reflectivity regions². On the other hand, this is the only estimator that enables us to obtain a precise Gamma law for the reflectivity estimate over a homogenous region.

6.2.2.2. Spatial filters

The reasoning we presented in the multi-look case above may be extended to the spatial case. The different looks will then come from neighboring pixel values. For complex images, we have again the spatial whitening filter, with vector \mathbf{Z} being constructed on a small window of values around the considered pixel. Likewise, in intensity we obtain the mean filter on a sliding window. Note that the processed image may well be a multi-look image; if so, we use relation [5.5] yielding the same result, which is the ML estimator if the pixel intensities are decorrelated. The mean filter, which was described as part of the standard image processing techniques, is the ML estimator (for decorrelated pixels) when applied to an intensity image. However, this is no longer true for amplitude images, since we would have to use squared amplitudes, according to the results established in the previous section.

Another spatial filtering which can be applied to radar images consists of averaging the logarithm of intensities and using the exponential function of the result. Such filters are not optimal in nature and do not correspond to the ML estimator. Reflectivity estimates are therefore highly biased. More importantly still, the estimator variance is not optimally reduced [ZRN 75, FJO 99a].

We will return in section 6.6 to how the window should be sized to calculate these spatial estimators.

² This is no longer true for targets for which the impulse response is heavily deteriorated [BRU 97].

6.3. Single-channel filters with *a priori* knowledge of the scene

In this section, we will deal with the case of a radar intensity image³ and study various reflectivity estimators. Depending on the complexity degree we choose, we will be able to consider speckle information (radar image formation process), or information on the underlying scene, or both simultaneously.

6.3.1. Linear MMSE filters

6.3.1.1. Lee and Kuan filters

We assume a stationary⁴, ergodic⁵ reflectivity of the scene, characterized by its variance and expectation (for their estimation, see section 6.6). Similarly, speckle S is characterized by its first order moments (its expectation is 1 and its variance $1/L$). We use the multiplying model written as (see section 5.3):

$$I = RS.$$

By applying the criteria (equations [6.1] and [6.2]) mentioned earlier relative to the linear MMSE, the estimator we obtain is:

$$\hat{R} = E(I) + k(I - E(I)) \quad \text{with } k = \frac{\text{var}(R)}{\text{var}(I)} \quad [6.3]$$

This filter has been proposed by Kuan [KUA 85] and is often named after him.

If the approximation $(R - E(R))(I - E(I)) \approx 0$ is used, we have the Lee filter [LEE 81], which is also widely used and very similar to the Kuan filter, except for its factor:

$$k = \frac{\text{var}(R)}{\text{var}(R) + \text{var}(S)E(R)^2}$$

³ For many images (e.g., the ERS radar precision image (PRI) products), the transmitted signal is in amplitude, not in intensity, and will have to be squared accordingly.

⁴ The process is stationary if its estimated spatial mean is invariant at any point in space.

⁵ The process is ergodic if its spatial average does not depend on the realization on which the estimate is based.

Between the variances of intensity, scene reflectivity and normalized speckle, we may note the following relation:

$$\text{var}(I) = \text{var}(R) \left(E(S)^2 + \text{var}(S) \right) + E(R)^2 \text{var}(S)$$

We then obtain k expressions in terms of variation γ_I and γ_S (see section 5.2.2.2):

$$k = \frac{1 - \frac{\gamma_S^2}{\gamma_I^2}}{1 + \frac{\gamma_S^2}{\gamma_I^2}} \quad \text{for the Kuan filter} \quad [6.4]$$

$$k = 1 - \frac{\gamma_S^2}{\gamma_I^2} \quad \text{for the Lee filter} \quad [6.5]$$

These two filters that are widely used in radar imagery behave alike for $L \geq 3$. On a very homogenous area, we have $\gamma_I \approx \gamma_S$, hence $k = 0$, which implies $\hat{R} = E(I)$. The value of the filter on a homogenous region is simply the local average. By contrast, on a very heterogenous region (e.g., close to a discontinuity), we have $\gamma_I \gg \gamma_S$, hence $k = 1$ for the Lee filter (and $\frac{1}{1+\gamma_S^2}$ for the Kuan filter). Thus, $\hat{R} = I$ (and, respectively, $\hat{R} = \frac{1}{1+\gamma_S^2} I + \frac{\gamma_S^2}{1+\gamma_S^2} E(I)$) and averaging is interrupted or limited on this sort of region where the best estimation of reflectivity is radiometry of the pixel itself.

On the other hand, it is preferable to impose $k = 0$ when the intensity empirical variation factor is lower than that of speckle, which would amount to squeezing k between 0 and 1. Regularization $k = 0$, which comes down to estimating R by the average, can and should be applied to any adaptive filter by using local statistical values wherever the value of the empirical variation factor nears that of the speckle coefficient [LOP 90b, LOP 93b].

6.3.1.2. Frost filter

This type of filter, which was proposed in [FRO 82], has the benefit of taking scene correlation into account explicitly. The previous assumptions still hold, but we

further assume that the autocorrelation of the scene is an exponential function decreasing with range d so that:

$$\rho_R(d) = e^{-ad}$$

As demonstrated in [MAI 77], this model often works fine for natural images.

A Wiener filter is the mean square error minimizing the estimator in this case [STE 67]. If speckle is regarded as white noise, the impulse response of the filter is an exponential function:

$$m(d) = \alpha e^{-\alpha d}$$

The filtered image is then obtained by convolution with this filter.

Parameter α appearing in the filtering formula is found from the relation:

$$\alpha^2 = a + 2a \frac{\gamma_I^2 - \gamma_S^2}{\gamma_S^2 (\gamma_I^2 + 1)}$$

Thus, all pixels take part in filtering, but their weight decreases on the one hand according to distance from the considered pixel and, on the other hand, to heterogeneity of the scene (under influence of term γ_I). Scene texture is taken into account through factor a . Still, the Frost filter is most often used in the form:

$$m(d) = K \alpha e^{-\alpha d} \quad \text{with } \alpha^2 = K' \gamma_I^2$$

K is a normalization constant aimed to avert mean value bias, is equal to the sum of the weights in the pixel's neighborhood. K' is another constant that helps adjust the amount of filtering: when K' is high, we have low filtering in all heterogeneity cases. By contrast, when it is low, the exponential function decrease is slower and filtering is more pronounced. Scene heterogeneity is taken into account anyway because of γ_I .

Several ways to improve this sort of filter, which is widely used, have been proposed [QUE 90]. Note that even where scene texture is introduced, speckle is still assumed to be decorrelated, which is rarely the case in intensity images (see section 5.3.1). On the other hand, speckle decorrelation can be achieved by resampling the SLC images, if available, after the spectrum has been whitened (see Chapter 5).

6.3.2. Non-linear MMSE filters

While the above analysis is based on the multiplying model, it makes no assumption as to the distribution of the scene. By contrast, the estimators we will now discuss require the knowledge not only of speckle distribution (Gamma and Goodman laws are often used, as seen in section 5.3.4.2), but also of the reflectivity distribution (Gamma distributions that help achieve good modeling of a wide range of natural textures are frequently preferred).

6.3.2.1. *T*-linear filters

The solution provided by these filters is close to the theoretical MMSE, i.e., to the *a posteriori* expectation, as it is no longer assumed to be linear, but rather “*t*-linear”, meaning that we assume a linear relation between R and a t power (not necessarily an integer) of I : $\hat{R} = a_t I^t + b$ [HIL 88, HIL 92]. The minimization of the MSE then leads to the following expression:

$$\hat{R} = a_t I^t + E(I) - a_t E(I^t) \quad \text{with } a_t = \frac{\text{cov}(I^t, R)}{\text{var}(I^t)}$$

Factor a_t has physical dimension I^{t-1} . Unlike in the linear $t = 1$ case, where the coefficient’s expression simply relied on the first and second order moments, its estimation is now based on the non-integer moments of intensity I and speckle S .

The generalized *t*-linear estimator is such that its value t_{opt} minimizes the MSE, for $0 \leq t \leq 1$. It is thus more general than its equivalent linear case (which we encounter for $t = 1$) and simplifies our approach to the true MMSE estimator which is the *a posteriori* expectation, which will be discussed further below.

We can show that the MSE in this case is written:

$$MSE(t) = var(R) - \frac{\text{cov}(I^t, R)}{var(I^t)}$$

Unfortunately, the optimal value of t is difficult to determine. In practice, we assume that we know the scene distribution and go on to estimate the parameters of this distribution enabling us to calculate $MSE(t)$, which we will then minimize with respect to t . Consequently, with a Gamma law for R , we will be able to perform the calculation and find optimal t . For $L = \alpha = 1$ and $L = 3, \alpha = 5$, the value that we find is close to 0.5. Unless we make an assumption on R distribution, our t values will be limited to 0.5 and 1.

6.3.2.2. APE filters

A posteriori expectation is the optimal solution as far as the least squares are concerned. It is obtained from relation:

$$\hat{R} = E[R|I] = \int_0^{+\infty} p(R)p(R|I)dR$$

The analytical solution is generally unknown, yet we can find its expression for some distributions of reflectivity R . One such case is where R follows a Gamma law. We then obtain [LOP 97a]:

$$\hat{R} = \sqrt{i_o} \frac{K_{\nu-L+1}(\sqrt{i_o})}{K_{\nu-L}(\sqrt{i_o})} \frac{E(R)}{2\nu} \quad \text{with } i_o = 4L\nu \frac{I}{E(R)}$$

where $\nu = \frac{1}{\gamma_R^2}$ is the shape parameter of the scene, $E(R)$ is its expectation (estimates of both are given in section 6.6) and finally L is the number of looks. This estimator is unbiased, yet harder to calculate because of Bessel function K arising in its expression, particularly when the value of ν is high. The t -linear estimator where $t = \frac{1}{2}$ is a satisfactory approximation of it. The value of factor α_t in this case is:

$$a_{\frac{1}{2}} = \frac{E\left(I^{\frac{3}{2}}\right)\left(\frac{L}{L+\frac{1}{2}}\right) - E\left(I^{\frac{1}{2}}\right)E(I)}{var\left(I^{\frac{1}{2}}\right)}$$

6.3.3. Non-linear MAP filters

We have focused on the mean square error criterion so far. However, the maximum *a posteriori* expectation is also widely used and, under some assumptions, can give some simple expressions for the estimators. *A posteriori* probability is written according to the Bayes rule:

$$p(R|I) = \frac{p(I|R)p(R)}{p(I)}$$

and we can find \hat{R} by maximizing this expression. Unless we have *a priori* knowledge of R , it will suffice to maximize $p(I|R)$, which is the ML estimator and yields in the pixel case $\hat{R} = 1$ (see section 6.2.2.1). Where there is *a priori* knowledge of $p(R)$, we calculate \hat{R} such that:

$$\frac{d \log p(I|R)}{dR} + \frac{d \log p(R)}{dR} = 0$$

Several estimators corresponding to different reflectivity distributions have been calculated [KUA 87, NEZ 92]. If we have a Gamma distribution for R , then R is obtained as the solution of second degree equation:

$$\hat{R} = \frac{1}{2\nu} \left(E(R)(\nu - L - 1) + \sqrt{E(R)^2(\nu - L - 1) + 4\nu L I E(R)} \right)$$

where again ν is the form scene parameter. Note that the MAP estimator, unlike the previous ones, is biased. The dynamics of reflectivity R are better preserved instead [LOP 97a].

6.4. Multi-channel filters

We have reviewed several filters that made it possible to restore the reflectivity of an intensity image. Most of them can be transposed to amplitude images, though their expressions may be more complex, given the shape of distribution. In this section, we no longer consider the case of a single datum, but rather one in which several channels are available to us. These “channels” may consist of very different data. We may have multi-temporal data, i.e., several intensity images acquired at different times, or polarimetric data, or “spatial” data where the neighborhood of a pixel is actually considered (see section 6.2.2.2). Whereas some general expressions given in this section (especially in section 6.4.2) may also apply to polarimetric data, a complete separate section (6.5) has been devoted to them.

6.4.1. Intensity data

We only consider intensity data in this section and therefore can apply optimal addition methods to approach the maximal number of looks NL , where N is the number of channels and L the original number of looks of every image.

6.4.1.1. Homogenous channels

I_i is the observed intensity on channel i and R_i, S_i is the corresponding reflectivity and speckle. We assume here that the channels are homogenous, i.e., on each of them we have constant local reflectivity (no texture in the scene). We then search for an unbiased estimation:

$$\hat{R}_i = \sum_{j=1}^N \alpha_j^i I_j \frac{E(I_i)}{E(I_j)}$$

Taking as a criterion the mean square error minimization, which also corresponds to minimizing the variance estimator in this scenario, we can show that, whatever channel is considered, the set of factors $(\alpha_i)_{1 \leq i \leq N}$ is the solution to the equation system:

$$\sum_{i=1}^N \alpha_i = 1$$

$$\sum_{i=1}^N \alpha_i (\rho(I_i, I_j) - \rho(I_i, I_1)) = 0 \quad \forall j \in 2, \dots, N$$

If we denote by $p(I_i, I_j)$, the correlation coefficient between channels i and j :

$$\rho(I_i, I_j) = \frac{E[(I_i - E(I_i))(I_j - E(I_j))]}{\sigma_{I_i} \sigma_{I_j}}$$

This set of factors is therefore optimal for every channel. It is sufficient to simply calculate it for one channel and deduce the others by applying a radiometric correction such that $E(\hat{R}_i) = E(I_i)$ leading to relations $\hat{R}_i = \hat{R}_j \frac{E(I_i)}{E(I_j)}$. When all channel intercorrelations are zero (meaning the channels are entirely decorrelated in terms of intensity), we again run across the channel-averaging formula based on radiometric equalization.

The reflectivity variation factor estimated on homogenous regions is given by:

$$\gamma_{\hat{R}_i} = \frac{1}{\sqrt{L}} \sum_{j=1}^N \alpha_j^2 + 2 \sum_{k < l} \alpha_k \alpha_l \rho(I_k, I_l)$$

where L is the initial number of looks on the channels. The equivalent number of looks (hence the speckle reduction) is maximum when all channels are decorrelated.

When the channels correspond to a physically homogenous, speckled area, the correlation is weak and the filter is practically equal to the mean. Otherwise, i.e., if structures are present, the correlation rises and filtering is limited. Note that this approach can be generalized to complex cases [BRU 97].

6.4.1.2. Textured channels

The above analysis can also be conducted so that texture is taken into account on the different channels (the MSE minimum and minimal variance estimator no longer coincide here, as they did in the case described in section 6.4.1.1). We therefore obtain more complex expressions for the factors that may involve one set of factors per channel [BRU 94].

For the minimal variance estimator, we obtain the following equations:

$$\sum_{i=1}^N \alpha_i = 1$$

$$\sum_{i=1}^N \alpha_i \left(\rho(I_i, I_j) \gamma_{I_i} \gamma_{I_j} - \rho(I_i, I_1) \gamma_{I_i} \gamma_{I_1} \right) = 0 \quad \forall j \in 2, \dots, N$$

where coefficients of variation of the different channels are involved. Here again, we have a single set of coefficients, with N channels to be obtained by radiometric correction later on.

In addition, a multi-temporal merger method has been proposed to integrate empirically a number of radiometric and textural criteria as well as channel correlation [PEN 94].

6.4.2. A general case: the linear vector MMSE

This filter is designed to take into account the correlation between the different channels, which may be due to either scene or speckle correlation. However, this is a very general case, as textures on each channel may be different from one another.

The modeling we adopted goes as follows. \mathbf{Y} is the data vector on which we will refrain from making any assumption for now (we may have intensity or complex data corresponding to different dates, frequencies and polarizations). These different cases will be presented in detail further below. \mathbf{X} is the data vector we are trying to estimate (e.g., the vector of reflectivities in every channel) and V is a diagonal matrix connecting \mathbf{X} and \mathbf{Y} :

$$\mathbf{Y} = V\mathbf{X}$$

The estimator minimizing the MSE under constraint of linearity is:

$$\hat{\mathbf{X}} = E(\mathbf{X}) + \mathbf{C}_X E(V)^* \mathbf{C}_Y^{-1} (\mathbf{Y} - E(\mathbf{Y})) \quad [6.6]$$

where \mathbf{C}_X and \mathbf{C}_Y are the covariance matrices of \mathbf{X} and \mathbf{Y} , respectively. The main difficulty with this type of filter refers to estimating the covariance matrix \mathbf{X} while knowing the empirical covariance matrix calculated from observations \mathbf{Y} .

In the case of intensity data, expression [6.6] is considerably simplified as:

$$\mathbf{I}_i = R_i S_i$$

for every component of vector $\mathbf{Y} = \mathbf{I}$ (where $\mathbf{X} = (R_i)_{1 \leq i \leq n}$). Therefore, matrix V is simply a diagonal matrix consisting of normalized speckle for the diagonal terms, and we have $E(V) = I_d$ (identity matrix). The expression is simplified to:

$$(\hat{R}_i) = E(\mathbf{I}) + \mathbf{C}_{R_i} \mathbf{C}_{\mathbf{I}}^{-1} (\mathbf{I} - E(\mathbf{I}))$$

This filter is a multi-dimensional generalization of the Kuan filter.

The problem of estimating different statistical quantities, particularly those of the covariance matrices of observations and reflectivity, is addressed in section 6.6.

The vectorial LMMSE solution for intensity data may be used in several types of situations:

- case of multi-date data: when several images or looks captured at different moments are available and make up vector \mathbf{I} . The channels may be generally correlated depending on the acquisition conditions and this information is taken into account by the covariance matrix;

- case of single-date piece of data: a vector can be formed based on a central pixel and its closest neighbors; we speak of spatial vectorial estimation in this case.

Note that this is one of very few techniques that make it possible to take into account the scene and speckle spatial correlation with different textures in the channels.

6.5. Polarimetric data filtering

According to equation [5.18] (product model), the observations come from the product between a multi-variate centered Gaussian vectorial random variable (of dimension 3 or 4) \vec{G} and a scalar random variable \sqrt{v} representing texture. In contrast to the one-dimensional case where we could decompose the received signal into a product of “information” (reflectivity R) by “noise” (speckle n of average equal to 1), here we can no longer normalize \vec{G} to separate information from noise (except in the trivial decorrelation of polarimetric \vec{G} components). Due to the possibility of considering the different information to be filtered, two types of polarimetric filters with radically different objectives have been proposed:

- speckle reduction filters based on a combination of polarimetric channels (some of these filters have been outlined in previous sections); and
- polarimetric data restoration filters, i.e., estimations of covariance matrix C .

6.5.1. Speckle reduction filters based on a combination of polarimetric channels

Combining the polarimetric channels according to a Polarimetric Whitening Filter (PWF) has been proposed [NOV 90] as a speckle reduction method. The PWF via a linear combination transforms the channels and decorrelates them to ensure an optimal speckle reduction. Here, the optimality criterion is the minimization of the coefficient of variation (ratio of the standard deviation to the mean).

If we denote by P the transition matrix from \vec{S} to transformed vector (decorrelated channels) $\vec{w} : \vec{w} = P \cdot \vec{S}$ and by y the output value $y = \vec{w}^{*t} : \vec{w}$ we have:

$$\begin{aligned} E(y) &= E(\vec{w}^{*t} \cdot P^{*t} \cdot P \cdot \vec{S}) = E(\vec{S}^{*t} \cdot A \cdot \vec{S}) \\ &= E\left(\text{tr}(A \cdot \vec{S}^{*t} \cdot \vec{S})\right) = \text{tr}(A \cdot C) = \sum_{i=1}^d \lambda_i \end{aligned} \quad [6.7]$$

where $\mathbf{A} = \mathbf{P}^{*t} \cdot \mathbf{P}$ is Hermitian, d is equal to 3 or 4 and λ_i are eigenvalues of $\mathbf{A} \cdot \mathbf{C}$. Likewise, using equation [5.16], we find:

$$\begin{aligned} E(y^2) &= E\left(\left[tr(\mathbf{A} \cdot \bar{\mathbf{S}}^{*t} \cdot \bar{\mathbf{S}})\right]^2\right) = E\left(\left(\sum_{i=1}^d \sum_{j=1}^d A_{ij} \cdot S_i^* \cdot S_j\right)^2\right) \\ &= \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d A_{ij} \cdot A_{kl} \cdot (C_{ji} \cdot C_{lk} + C_{li} \cdot C_{kj}) \\ &= \sum_{i=1}^d \sum_{j=1}^d A_{ij} \cdot C_{ji} \cdot \sum_{k=1}^d \sum_{l=1}^d A_{kl} \cdot C_{lk} + \sum_{i=1}^d \sum_{j=1}^d \sum_{k=1}^d \sum_{l=1}^d A_{ij} \cdot A_{kl} \cdot C_{li} \cdot C_{kj} \end{aligned}$$

i.e.:

$$E(y^2) = [tr(\mathbf{A} \cdot \mathbf{C})]^2 + tr((\mathbf{A} \cdot \mathbf{C})^2)$$

therefore:

$$\text{var}(y) = tr((\mathbf{A} \cdot \mathbf{C})^2) = \sum_{i=1}^d \lambda_i^2 \quad [6.8]$$

The minimization of the coefficient of variation that is equal to $\sqrt{\sum_{i=1}^d \lambda_i^2} / \sum_{i=1}^d \lambda_i$, is obtained when all eigenvalues are equal, i.e.:

$$\mathbf{A} \cdot \mathbf{C} = \alpha \cdot \mathbf{Id}_d$$

where α is any scalar and \mathbf{Id}_d is the identity matrix of dimension d . Finally, the coefficient of variation reaches its minimal value $1/\sqrt{d}$ ($= -4.8$ dB, for $d = 3$) if and only if $\mathbf{A} = \mathbf{C}^{-1}$ with the exception of a constant.

In summary, when data are single-look matrices \mathbf{S} , the filtered image is created by $y = \tilde{\mathbf{S}}^* \cdot \tilde{\mathbf{C}}^{-1} \cdot \tilde{\mathbf{S}}$, where $\tilde{\mathbf{C}}$ is the covariance matrix locally estimated through a window centered at the pixel involved (adaptive version of the PWF). In the case of multi-look data, either in the form of a covariance or Mueller matrix, we will take $y = \text{tr}(\tilde{\mathbf{C}}^{-1} \cdot \Sigma)$ where $\tilde{\mathbf{C}}$ has the same meaning as above and Σ is the data covariance matrix at the pixel involved.

NOTES:

- the fact that $\tilde{\mathbf{C}}$ is estimated using a window introduces an additional spatial correlation among pixels;
- assuming reciprocity ($S_{hv} = S_{vh}$) and the decorrelation of co- and cross-polar channels (across-track symmetry), the expression of y is:

$$\frac{|S_{hh}|^2}{\tilde{\sigma}_{hh} \left[1 - \frac{|\tilde{\rho}|^2}{\tilde{\sigma}_{hh} \cdot \tilde{\sigma}_{vv}} \right]} + \frac{|S_{hv}|^2}{\tilde{\sigma}_{hv}} + \frac{|S_{vv}|^2}{\tilde{\sigma}_{vv} \left[1 - \frac{|\tilde{\rho}|^2}{\tilde{\sigma}_{hh} \cdot \tilde{\sigma}_{vv}} \right]} - \frac{2 \cdot \text{Re}[S_{hh} \cdot S_{vv}^* \cdot \tilde{\rho}^*]}{\tilde{\sigma}_{hh} \cdot \tilde{\sigma}_{vv} \cdot \left[1 - \frac{|\tilde{\rho}|^2}{\tilde{\sigma}_{hh} \cdot \tilde{\sigma}_{vv}} \right]} \quad [6.9]$$

where $\tilde{\sigma}_{ij} = E(|S_{ij}|^2)$ (with i, j having the value h or v) and $\tilde{\rho} = E(S_{hh} \cdot S_{vv}^*)$;

- in the general case of the K distribution (taking texture into account), the coefficient of variation is written [NOV 90]:

$$\sqrt{\frac{\nu+1}{\nu} \times \left[\sum_{i=1}^d \lambda_i^2 \Big/ \left(\sum_{i=1}^d \lambda_i \right)^2 \right] + \frac{1}{\nu}}$$

The minimal value reached (again for equal eigenvalues) is then:

$$\sqrt{\frac{1}{\nu} + \frac{1}{d} \times \frac{\nu+1}{\nu}}$$

A variant based on intensity images only

The filter proposed in [NOV 90] only combines intensities measured at the different polarizations, $|S_{hh}|^2$, $|S_{hv}|^2$ and $|S_{vv}|^2$. The resulting expression introduces $|\rho|^2$, which has to be estimated by, for example, the Siegert relationship. At the same time, the expression of y in the general case where the three intensities have any correlation was established in [LEE 91]. The authors argued that instead of one image of y , the filter output should consist of three different images y_{hh} , y_{hv} and y_{vv} , corresponding respectively to intensities $|S_{hh}|^2$, $|S_{hv}|^2$ and $|S_{vv}|^2$. To this end, they simply normalized y so that the local mean of the considered channel is observed. Minimizing the coefficient of variation is then equivalent to minimizing variance.

The advantage of this filter is that it can be generalized to any multi-channel data, not necessarily polarimetric data, e.g., multi-look or multi-temporal data as in [BRU 97].

Figure 6.1a shows the result of applying the Novak and Burl PWF to AirSAR data acquired at L-band over the Orgeval site in France. Figure 6.1b presents the difference between this image and the simplified version based on intensities alone. The differences are mainly found at low incidences or on some agricultural plots.

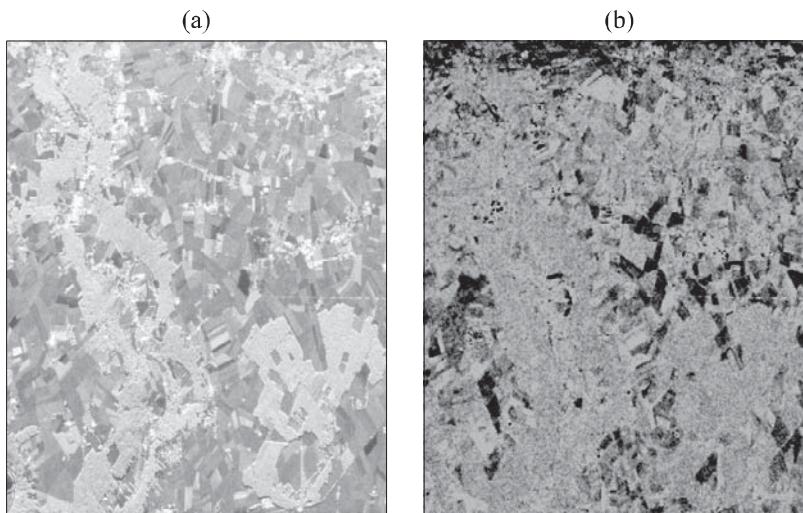


Figure 6.1. Novak and Burl PWF: (a) output image at L-band (site: Orgeval, France); (b) difference from the variant solely based on intensity images

6.5.2. Polarimetric data restoration filters

In the previous section, filters were based on the use of different polarimetric data (to obtain one or more images) with minimum coefficient of variation. Conversely, here it is the polarimetric information itself that we are now trying to retrieve. We will therefore seek to entirely restore matrices \mathbf{S} or \mathbf{C} . The filters to retrieve polarimetric data rely on the assumed distributions of \mathbf{S} and \mathbf{C} . Maximum likelihood is the optimality criterion used in this case.

The homogenous area case

In the case of homogenous areas, $\bar{\mathbf{S}}$ has a Gaussian distribution and the ML estimator of \mathbf{C} is the mean over the considered region [GOO 75]:

$$\tilde{\mathbf{C}} = \frac{1}{n} \cdot \sum_{k=1}^n \bar{\mathbf{S}}_{(k)} \cdot \bar{\mathbf{S}}_{(k)}^* \quad [6.10]$$

In the case of multi-look data or data in the form of Mueller or covariance matrices, we denote by $\Sigma_{(k)}$ the covariance matrix observed at pixel k . Σ follows a Wishart distribution (see section 5.3.5), and the ML estimator is once again the result of a mere averaging of observations:

$$\tilde{\mathbf{C}} = \frac{1}{n} \cdot \sum_{k=1}^n \Sigma_{(k)} \quad [6.11]$$

The heterogenous area case (product model)

In the case of textured areas, the covariance matrix $\Sigma_{(k)}$ observed at pixel k corresponds to the product of random texture variable μ_k by a covariance matrix $\Sigma_{0(k)}$ having a Wishart distribution with mean value \mathbf{C} ($\Sigma_{0(k)}$ is the covariance matrix which, without texture, would have been observed and its variations around \mathbf{C} are due to speckle). The ML estimator of \mathbf{C} is then:

$$\mathbf{C} = \frac{1}{n} \cdot \sum_{k=1}^n \Sigma_{0(k)} = \frac{1}{n} \cdot \sum_{k=1}^n \frac{1}{\mu_k} \Sigma_{(k)} \quad [6.12]$$

As far as textured regions are concerned, the problem is to estimate μ_k and \mathbf{C} at the same time, as on the one hand $\tilde{\mathbf{C}}$ is estimated according to μ_k (equation [6.12]) and, on the other hand μ estimation requires the knowledge of \mathbf{C} . The ML estimator, for example, is:

$$\mu = \frac{1}{d} \text{tr}(\tilde{\mathbf{C}}^{-1} \Sigma) \quad [6.13]$$

where μ and Σ are the values of texture and of the covariance matrix observed at pixel k . With the exception of factor $1/d$, we recognize here the expression of the PWF presented in the previous section. Other estimators such as MMSE are also used.

Séry then proposed an iterative estimation of μ and Σ : assuming a homogenous region, a first estimation of \mathbf{C} ; texture μ_k is estimated at every point of the working window; it is thus possible to re-estimate \mathbf{C} (according to equation [6.12]); and so on up to empirical convergence [SER 97].

Figure 6.2a shows the phase difference of co-polar channels after filtering. Comparing the phase difference images after filtering and before filtering (Figure 5.20a), we note that noise has been reduced quite considerably. Figure 6.2b shows the texture parameter μ obtained during the same filtering process.

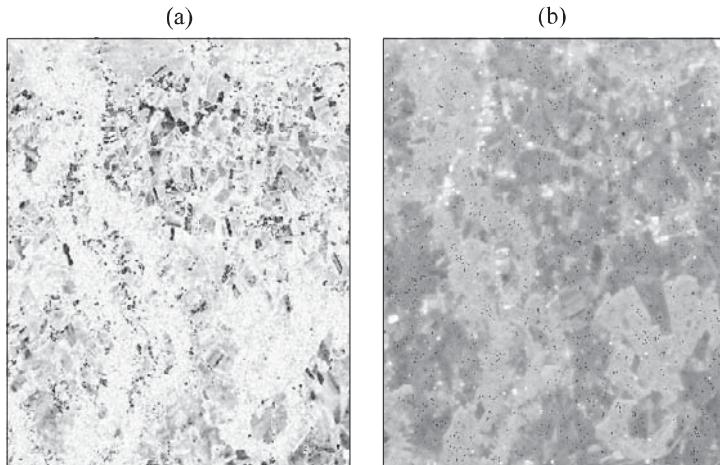


Figure 6.2. Polarimetric information filtering by the Séry method: (a) post-filtering image of the phase difference of co-polar channels hh and vv at L-band (Orgeval, France); (b) image of the corresponding texture parameter (ML estimator)

6.6. Estimation of filter parameters

Every estimator we have reviewed so far involves a number of statistical quantities (e.g., expectation, variation factor, covariance matrix) relative to observations, speckle and reflectivity. Their values necessarily impact on the quality of filtering, and their calculation gives rise to problems. On the one hand, since all we have available is observations, we must be able to calculate the statistical quantities related to normalized speckle and especially those related to reflectivity; this aspect is discussed in the next section. On the other hand, a reliable calculation of these estimators requires a large number of samples, which affects the accuracy of the filters (section 6.6.2).

6.6.1. *Formulae for statistical estimation*

All statistics (expectation, standard error, covariance matrix) related to observations are empirically calculated by assuming a locally stationary scene. A region is made of pixels of intensities I_1, \dots, I_N ; we then have:

$$\widehat{E(I)} = \frac{1}{N} \sum_{i=1}^N I_i$$

$$\widehat{\sigma_I}^2 = \frac{1}{N-1} \sum_{i=1}^N (I_i - \widehat{E(I)})^2$$

Speckle statistics may be provided by the sensor's physical parameters (number of looks) or calculated over a particularly homogenous test region (no recalculation will be needed to process images from the same sensor in this case).

Those from scene R are given by the following formulae based on intensity statistics:

$$E(R) = E(I)$$

$$\gamma_R = \sqrt{\frac{\gamma_I^2 - \gamma_S^2}{1 + \gamma_S^2}}$$

The reflectivity form parameter is given by $\nu = \frac{1}{\gamma_R^2}$.

In the case where R is distributed according to the Gamma law, ML estimators of its parameters have been proposed in [OLI93] and [JOU 93]. Comparisons with other estimators can be found in [BLA 94] and [ISK 99].

As for the covariance matrix of reflectivity in the case of intensity data, it can simply be deduced from the speckle and intensity covariances:

$$\rho(R_i, R_j) = \frac{1}{1 + \rho(S_i, S_j)} \left[\rho(I_i, I_j) - \rho(S_i, S_j) E(I_i) E(I_j) \right]$$

6.6.2. Choice of a calculation window

The shape and size of the statistical calculation window are key parameters in every one of these filters. They also have to meet two opposite requirements. On the one hand, the calculation has to be made on pixels sharing the same physical region (field, forest, urban area, etc.) as the considered pixel. On the other hand, the number of pixels used in the calculation must be as large as possible. Indeed, the number of independent samples that is required for a reliable statistical evaluation is in the order of 300 [LAU 89, GOU 96]. In practical terms, the deduced windows are prohibitively large and pixels from other regions have to be taken into consideration if the windows are square-shaped.

Therefore, Lee [LEE 81] and Kuan [KUA 85] proposed using edge detectors to locally check a processing window for homogeneity, but this would only meet the first constraint.

On the other hand, Wu and Maître [WU 92] proposed a solution that helps take both sides of the problem into account. It consists of looking in a pixel to find the largest homogenous window to which it belongs. The method is a region growth technique, starting from a square shape. The variation factor of variation Δy during growth makes it possible to distinguish among different cases and guide our enlargement of the window towards the most homogenous area available. The drawback of this approach is that, while filtering is overall more vigorous and preserves the edges well, the output image tends to reproduce the regions that were identified during growth and the restored image has a “pointillist” appearance as a result.

In every case, statistical calculations are performed on the largest possible window around the pixel. Proper filtering is then carried out with a Kuan filter, though other filters may do just as well.

Another way to resolve the problem consists of selecting large windows (e.g., 15x15), but pairing edge, line and specular reflector detection to the filtering in order to preserve them. This is the approach adopted by Lopes *et al.*, who used a Gamma-MAP filter and performed these detections even as filtering was in progress [LOP 93b]. These processes (edge, line and target detections) are the subject of Chapter 8. Such parallel detection of elements within the window of analysis helps to improve performance in all the estimators we reviewed above.

Finally, it might be quicker to segment the image beforehand and calculate the parameters either in the entire region, or at the intersection of a large window and the region to which the considered pixel belongs [FJO 97], but this raises a segmentation problem.

6.7. Filter specificities

6.7.1. Statistical evaluation of the various estimators

As far as conditional bias is concerned (i.e., the comparison of R and expectation of R for $p(I|R)$), it is hard to provide general performances of the different estimators. They are a function of the relative value of R in terms of its own mean

$$\mu = \frac{R}{E(R)} \quad (\text{texture parameter})$$

Note, however, that we tend to underrate reflectivity when $\mu > 1$ and overrate it when the opposite is true (in summary, the estimated values are statistically closer to the mean value than they should). The MAP estimator stands out among all others by its better performance in terms of conditional bias and MSE when $\mu \ll 1$. Conversely, its performance is less than those of other estimators when L is low and heterogeneity is high. Single-look images are a litmus test for the differences between the various filters. In [SER 97] detailed results are given for a Gamma distribution of reflectivity when its parameters are perfectly well known; also given are curves showing a slightly decreased performance of estimators based on this assumption when the texture distribution no longer observes this law.

As for the overall bias, APE and MSE estimators are unbiased in this case, unlike the MAP estimator which systematically underrates average reflectivity $E(R)$ in cases where the *a posteriori* distribution is particularly asymmetric with a mode

clearly lower than the distribution mean. Therefore, filters differ most from one another where we have low values of L and strong heterogenities.

The first type of bias (or conditional bias) will influence speckle-related texture reconstruction errors, whereas the second shows errors in estimating mean radiometry $E(R)$. The value of the latter bias will therefore matter if the user extracts quantitative data from physical quantities, specifically from the average backscattering radar factor of a surface in this case. To quantify the effect of the first type of bias on segmentation or classification methods performance using texture on filtered images, we will need to examine yet more (conditional and overall) criteria, including variance and the mean square error incorporating bias effect.

Note that for overall quantities and MSE especially, the APE estimator is the best point estimator in terms of MSE, irrespective of the distribution of R . The relationship between its MSE and variance is:

$$MSE + \text{variance} = \text{var}(R)$$

This proves its variance is systematically lower than that of texture R , which is a characteristic that will appear even more marked in other estimators. In the case of a Gamma distribution where parameters are perfectly well known, we find that the normalized variance of all estimators (used instead of variance to correct potential bias effects) is also systematically less than the normalized real variance of R . This means that point estimators do not preserve the texture heterogeneity and reduce the reflectivity deviation dynamic. Actually, as we have pointed out, they do so all the more as we have lower L and higher heterogeneity. On average they tend to oversmooth the texture in the presence of heterogeneity, although the opposite would actually be expected. The MAP estimator is still the best as far as this criterion is concerned. Beyond a dozen looks, all estimators are virtually equal to one another, but fall short of the mark unless the number of looks is higher [LOP 97a, SER 97].

Point estimators such as the Frost filter and linear vector MSE that take spatial correlation into account here yield what are in principle the best results (lower MSE and variance closer to the real value). In practice, though, these effects may be altered by estimation problems related to the unknown parameters characterizing R statistics. At that point, the quality of the available estimators, generally those based on the method of moments and less frequently on the ML method, comes into play. In the case of adaptive (contextual) filters, the results very much depend on the size of the analysis window, provided that we carried out the check for non-stationarity, boundaries and strong scatterers within it. Concerning the latter, a different processing should be applied that would consist of either keeping the initial values

unchanged, or deconvoluting to locate the target more accurately and estimate its RCS by integrating the energy of its response which is assigned to the pixel in which it lies.

6.7.2. Results on an ERS-1 image

The performances of various filters on a single-look image can be evaluated visually in Figure 6.3. The results, as explained earlier, are definitely more conclusive when structure detection is incorporated in the filtering procedure (see results f) to h)). The result of a linear vector MMSE filter (see section 6.3.2) on a temporal series of 6 ERS images is given in Chapter 7, complete with a calculation of local statistics on regions obtained by image segmentation and assuming a temporal decorrelation of the speckle in estimating covariance matrix M . For classification purposes, associating this vector estimator with segmentation leads to a considerable improvement of results compared with those provided by the other filters reviewed in this chapter.

If we numerically compare the performances of different filtering algorithms, considering the chain in its entirety, we find that the results in general are neither comparable nor stable (“clashing” would be a more appropriate term), depending on whether estimators or detection operators have been implemented or not, and if they have, on their effectiveness. The results will also depend on the criteria used with respect to some objective measures on the image (bias, MSE, edge smoothing), which do not take values unless the image parameters are perfectly well known, i.e., unless we deal with a synthesized image. The characteristics of reflectivity and speckle must be precisely simulated in this case, in order that we get as close as we can to the real images, while taking into account basic characteristics such as spectrum and sampling from which spatial correlation, statistical distribution, added thermal noise, etc., are derived. Another way to evaluate the segmentation methods and filters was proposed in [CAV 98]. It consists of analyzing image statistics of the ratio between unfiltered and filtered images, which ideally are those of pure speckle. Still, we must be careful with the overall numerical results, since speckle is not necessarily present everywhere.

Other criteria are directly related to the application based on the filtered image (e.g., classification, see Chapter 7, detection and photo interpretation, extraction of physical parameters, etc.). Generally speaking, all studies agreed to say that the performance of a filter depends on the intended application [DUR 87, NAJ 97, XIE 99, FOU 99] and that there is no such thing as a universal filter yet.

6.8. Conclusion

The purpose of this chapter was to present the various filtering methods used in radar imaging. We focused on estimators and *a priori* information on the scene that are most commonly used. There are of course other methods that use, for example, wavelet transforms [FOU 96, BIJ 97, DON 98, HER 98], or different *a priori* knowledge of the scene, such as Markovian models [WHI 94, TUP 97, DAT 98, WAL 99], or that rely on logarithmic transforms of the image via homomorphic filters [ARS 84, FRA 95].

Comparing all these filters and all their variants is no easy task in practice (see for instance [LEE 94c]). Their performances most sharply differ when single-look data are at stake, but are equivalent to one another once we go beyond a small number of looks. In fact, choosing a particular filter essentially depends on what the user plans to do with it and is often an empirical decision. It may appear, for example, more interesting to apply a Lee or Kuan filter rather than averaging if what we want to use is wavelet transforms. A Frost filter that takes into account spatial correlations will preserve structures better. A more sophisticated filter using statistical models [DAT 98] seems to work better on high resolution data. Unfortunately, although comparisons are regularly made, no general rule seems to emerge.

To conclude, we can say that at present neither the filters nor the estimators are up to task. Indeed, the visual improvement that has been obtained is not always good enough for the photo interpreters which sometimes lose track of their marks on speckled images. The same applies to automatic processing: pre-filtered images are not always the best solution and sometimes it would be preferable to work on raw images (e.g., when we want to detect fine structures such as pixel targets and lines). By adapting detection methods to radar image properties, we avoid the need for intermediate stages. In these circumstances, the methods that best estimate local surface properties (e.g., reflectivity, texture parameters, polarimetric and interferometric parameters, etc.) and detect potential deviations of the data with respect to speckle conditions as well as borders between regions, can easily be substituted for one another.

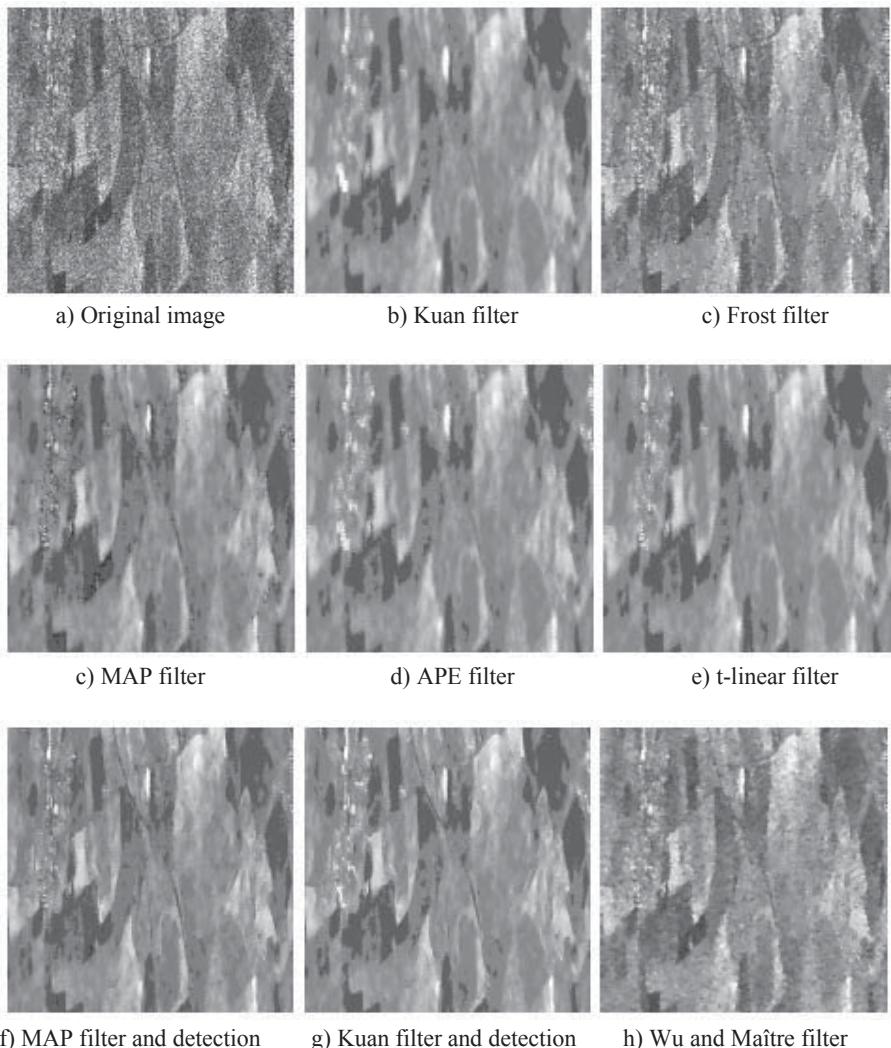


Figure 6.3. Results obtained with different filters: a) original single-look image © ERS-1; b) to e) results of filtering as provided by different point filters (statistics are estimated over square windows, 7x7 in size in every pixel; f) and g) MAP and Kuan filters with statistical estimations using the detection of structures, edges, lines and bright spots; h) Kuan filter with a search for the largest homogenous region according to the Wu and Maître method

This page intentionally left blank

Chapter 7

Classification of SAR Images

Detecting primitives in a scene (see Chapter 8) is a major stage in automatic image processing. Classification, the purpose of which is to identify land cover, is another important operation. Unlike detection, which is often local in scope, classification is a regional, even global, approach to the scene involved. Its objective is to assign to every pixel in the scene a label that will bring together all pixels similar in nature and identify the class to which that pixel belongs.

Classification methods are the foundation on which shape recognition techniques are based [DUB 90]. They rely on a variety of tools provided by data analysis, discriminant analysis, probabilistic approach, decision theory and many other fields of applied mathematics. Not all of these techniques can be used in radar imaging. In addition, the remote sensing community has long been using optical field imaging tools. Therefore, it is naturally this type of methods that have predominantly been applied and extended to the radar domain. We will skip a general classification techniques presentation and instead describe only the methods that have been explored over the past few years and whose performances were good enough to have them rated as applicable to radar imaging. They mainly include Bayesian methods good at exploiting our *a priori* knowledge of the radar signal distribution laws. By contrast, other methods such as neural networks [HAR 94, CHE 96] or projection methods (main component analysis, factor analysis, etc.), albeit well developed in other shape recognition areas, have scarcely been used in radar imaging so far.

Classification methods traditionally fall into two categories:

- *unsupervised classifications*, which have no use for any *a priori* knowledge of data; they rely on the discriminatory features of data (radar signal in our case) to separate the pixels into different classes as homogenous as possible. The number of these classes is often unknown and has to be determined by the classifier. The methods put in place are often iterative and build the classes gradually starting from optimized distances or pseudo-distances (i.e., measures that do not verify all axioms of distance);
- *supervised classifications*, which rely on *a priori* knowledge of some information about the classes we are trying to identify, e.g., the nature of these classes (forests, crops, urban areas), their specific properties such as the probability laws of their attributes (radiometry, texture, polarimetric or interferometric properties, etc.).

In the first case, classification needs to be interpreted, so the next thing we have to do is to assign semantics to the classes we have found. In the second case, classification will require previous training so that we may adapt our knowledge to the scene, the sensor or to particular conditions of the current experiment. Despite these differences, classification methods, whether supervised or not, share several principles discussed below.

7.1. Notations

$\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ is the finite set of cardinal n classes and d is the vector dimension of y characteristics assigned to every pixel of the image and x the class to which the pixel belongs. In a scalar image, y is the amplitude or intensity, as the case may be, and $d = 1$. In a polarimetric image, d will be 3 or 4 (see section 4.1.5). In a temporal series, d is the number of dates. The N pixels to be classified are called *observations* or individuals. Y represents the field of N measurements and X the field of classes.

In remote sensing, the supervised techniques rely on a manual assignment of the training plots; these are perfectly well known terrain areas identified on the images and constitute what we call the “ground-truths”. Part of these samples is used to learn the classes and the rest of them to verify classification results.

Confusion matrix M_c is an $n \times n$ matrix, which is calculated from reference samples. Diagonal elements $M_c(i, i)$ represent the number of pixels in class i that have effectively been assigned to it as a result of classification. Non-diagonal element $M_c(i, j)$ is the number of i -class pixels that have been assigned to class j . By dividing $M_c(i, j)$ by the number of pixels of i -class samples (denoted by m_i), we

obtain the percentage of i -class pixels classified in j (denoted by $P_c(i, j)$). The following measures are calculated from these matrices:

- the rate of i -class pixels correctly classified, i.e. $P_c(i, i) = M_c(i, i)/m_i$, also known as user accuracy or quality;
- the “commission” error defined as 100 minus user quality (expressed as a percentage);
- the producer i -class quality or accuracy, i.e. $P_{Pc}(i, i) = M_c(i, i)/m_{.i}$, where $m_{.i}$ is the number of pixels classified in i (sum of rows in column i); a high rate shows this class has rarely been assigned to samples from other classes. Conversely, a low rate indicates this class has often been selected instead of others;
- the “omission” error defined as 100 minus producer quality (expressed as a percentage);
- the overall accuracy (percent correctly classified – PCC):

$$PCC = \frac{1}{n} \sum_{i=1}^n P_{Uc}(i, i) = \frac{1}{n} \sum_{i=1}^n M_c(i, i) / m_i;$$

- the Kappa factor [COH 60] given by:

$$\kappa = \frac{N_e \sum_{i=1}^n M_c(i, i) - \sum_{i=1}^n m_i \cdot m_{.i}}{N_e^2 - \sum_{i=1}^n m_i \cdot m_{.i}}$$

where N_e is the total number of sample pixels. This factor takes into consideration every confusion matrix element. Besides, a Kappa factor can be given for every class.

7.2. Bayesian methods applied to scalar images

We will confine our presentation to the Bayesian methods used for SAR image classification. It is indeed important, as we have seen in Chapter 5, that speckle be well modeled statistically. This type of method that can make explicit use of our knowledge of the signal and noise probabilities is therefore most well suited for this purpose. We will identify *pixel methods* (see section 7.2.1) classifying every pixel

without any reference to the context, *contextual methods* (see section 7.2.2) that use the neighborhood of every pixel, and *global methods* (see section 7.2.3) using a model over the entire image.

Within the Bayesian representation, X is the random field of classes (a realization x of which is sought); Y is the field of observations (where a realization y is available) and $p(X, Y)$, $p(Y|X)$, $p(X|Y)$ are joint, conditional and *a posteriori* probabilities, respectively.

Probability $p(Y|X = x)$ is the probability that observed image y comes from a class x image. It is often called an *observation model*, as it depends on the system that makes the observation and can theoretically be determined by modeling the sensor response over a region. Indeed, the same x class scene may lead to different observed images depending on observation conditions. As seen in Chapter 5, the distribution of observations in a coherent system such as SAR is quite irregular, much more so than in an incoherent optical system. Within the framework of supervised classification, this term can be estimated either theoretically or empirically, on samples which have been selected in a supervised way.

Probability $p(X = x)$, called *a priori* probability of the label field, is the probability of a class image independently from observed y . Class images are little known *a priori*. Still, it is up to us to include in this term every piece of information we can find on the solution that we seek. If nothing at all is known of the solution, we choose a uniform distribution by default. We will see when we discuss global methods how the Markov field theory enables us to calculate the *a priori* probability of an entire image.

Based on these two probabilities, we can use the Bayes formula to calculate *a posteriori* probability $P(X|Y = y)$, i.e. the probability of an x class image if observation y is provided:

$$P(X|Y = y) = \frac{p(Y|X = x)p(X = x)}{p(Y)}.$$

With the knowledge of this probability, the last stage will consist of determining the x image that will be retained based on this distribution. As in Chapter 6, several criteria (maximum likelihood (ML), maximum *a posteriori* (MAP), etc.) can be used to select x once probability field $P(X|Y = y)$ is known.

Concerning the methods that will be described below, we will speak of supervised classification when parameters involved are learned using knowledge sets or empirically fixed by the user. We will call the classification unsupervised when the process incorporates an automatic learning of the parameters during classification¹.

7.2.1. Pixel-based Bayesian methods

The methods included in this family are aimed at classifying every pixel independently of its context, based on the sole observation of the pixel itself. This type of classification method is fundamental in data analysis where we cannot generally rely on strong dependencies among individuals to improve classification. Pixel methods are widely used in optical image processing because they provide correct results at a low calculation cost and can therefore be found in most image processing software packages.

7.2.1.1. Supervised methods

The first stage in supervised methods consists of determining conditional and *a priori* probabilities.

Distribution $p(y|X = \omega_k)$ is assimilated to the empirical distribution, as measured on real ground-truth samples corresponding to this k class, or to a smoothed version of this empirical distribution.

Another solution consists of selecting a system of parametric distributions, where parameters are adjusted to the collected samples (e.g., so that identical instants are obtained). While Gaussian laws are frequently used in incoherent imaging, in radar imaging we exploit the known speckle results. When we deal with intensity images of agricultural scenes, for instance, we can approach distribution $p(y|X = \omega_k)$ of each class with Gamma laws (see Chapter 5, equation [5.5]) with the same mean and standard deviation as the samples. A Gaussian distribution can be adopted, especially for multi-temporal data, if we have homogenous regions and a high number of looks. If we have few looks and regions ranging from homogenous to very heterogenous, the K or Pearson system distributions can be adopted (see section 5.4.1.3), but they will not allow us to accurately model correlated multi-temporal data beyond $d = 2$ channels as the multi-variate Gaussian law would enable us to do.

¹ It should be emphasized, however, that this distinction may be quite arbitrary, as training models often resort to quite substantial *a priori* knowledge, i.e., we will most of the time have some degree of supervision.

In polarimetry, however, when we have identical textures in channels, the K law makes it possible to consider the covariance channel matrix.

In the second stage, we assign each pixel to the class maximizing its observation probability. The Bayesian optimization is here confined to comparing n probabilities. If no *a priori* probabilities are available, we can use an ML classifier which maximizes $p(y|X = \omega_k)$, i.e. that assumes class equiprobability. If some classes are very rare, the MAP will tend to penalize them, so we would rather choose better suited criteria such as the Neyman-Pearson criterion providing the class with the highest correct classification rate for a fixed misclassification probability.

The quality of obtained classification depends on the class overlapping. In the case of raw SAR images, i.e., which are unfiltered and have a low number of looks, the dynamic leads to high class overlapping and the quality of results is generally poor (see Chapter 5). Pixel classification gives a very noisy, pointillist-like appearance and rather low spatial coherence to class images. Correct classification for an ERS, JERS or RADARSAT channel does not exceed 15-25% on average. To improve these results, it is recommended that images should be filtered using the techniques reviewed in Chapter 6. It might also be interesting to work on the logarithm of filtered amplitudes since the Gaussian approximation then becomes acceptable. In this case and for multi-temporal images (from at least 3 dates upward), the rate of pixels correctly classified ranges from 55-70%; it increases with the number of dates (see section 7.3) and depends on the type of landscape.

7.2.1.2. Unsupervised methods

In their unsupervised versions, the Bayesian pixel methods have to learn parameters occurring in the process during the classification stage itself (rather than during any previous learning stage). These parameters are like the *a priori* class probabilities and the observation law $p(Y|x)$ parameters mentioned above, e.g., the mean, form parameter or Pearson system parameters, according to the models that have been adopted. We will denote by θ the vector of these parameters. Several algorithms can be used for this estimation.

Among them is the estimation-maximization (EM) algorithm [DEM 77] that maximizes the likelihood of observation y with respect to model parameters θ and then re-estimates parameters θ , based on the classification thus obtained. These two stages are iterated until the likelihood converges to a peak, which unfortunately is local rather than absolute, and depends especially on the initial value of parameters θ_0 . The choice of a good initialization is crucial not only for the final classification quality, but also for the convergence rate.

On the other hand, EM requires that the number of classes should be known. If we do not know this number, we can still supply an upper bound to the algorithm, but some convergence problems may occur. To circumvent this shortcoming, we can use the stochastic version of the EM or SEM algorithm [DEL 97a], which helps deal with the fact that we only know the upper class number bound. SEM relies on the same stages as EM, but uses a simulation of the image conditionally classified to observed data y , by knowing the *a posteriori* probabilities $p(X = \omega_k|y)$ estimated at the previous iteration.

These methods have the advantage of being entirely automatic. Applied to SAR images, however, they yield poorer results than the supervised methods used in incoherent optical imaging.

Several examples of single-date and multi-date supervised Bayesian pixel classification are given in section 7.3. As mentioned earlier, one of the major drawbacks of this type of methods is the lack of spatial coherence of their results which leads to a very noisy appearance of the final classifications. The two families of methods that will be examined below make it possible to correct such shortcomings which are characteristic of pixel classification methods.

7.2.2. Contextual Bayesian methods

Pixel-based methods use a set of *a priori* probabilities that is constant across the entire image and thus does not depend on pixel context. However, the statistical properties of an observed scene are locally coherent and the interpretation of a pixel depends on its neighbors at least within some neighboring area, if not at the entire image scale.

Contextual classification methods make the most of the pixel's spatial context at decision stage. We can imagine two ways of looking at the pixel context: either we impose a constant context to each pixel, or we pick for each pixel those neighbors that are particularly similar to it. The latter solution implies that a classification should follow segmentation.

We will make a distinction here between classification and segmentation. The goal of classification, as we have seen above, is to identify on an image the classes each pixel is related to. As for segmentation, it is aimed at dividing images into contiguous homogenous regions. A very fine segmentation (i.e., according to a very strict homogeneity criterion) is a good preliminary for classification, as all pixels in the segmented area have radiometric, textural and other attributes quite close to one

another and thus are likely to belong to the same class. This will therefore provide us with a strategy frequently used to classify an image and consists of:

- performing a rather strict segmentation according to pixel similarity criteria or by edge detection [COC 96]; and
- assigning a class to every area identified through segmentation using the high number of measures associated with every pixel in the area.

The images of agricultural plots are homogenous fields generally well suited for segmentation-based classification.

7.2.2.1. Supervised methods

\mathbf{Y}_s is a random real Gaussian variable with dimension d , mean vector \mathbf{M} and covariance matrix \mathbf{C} . If \mathbf{M}_e and \mathbf{C}_e are, respectively, the empirical mean vector and covariance matrix of a certain number N_e of realizations of Y_s , the joint *a posteriori* distribution of \mathbf{M} and \mathbf{C} , conditioned on M_e , C_e and N_e , is the real Gauss-Wishart distribution [FU 76]:

$$\begin{aligned} p(\mathbf{M}, \mathbf{C} | \mathbf{M}_e, \mathbf{C}_e, N_e) = & \frac{\sqrt{|N_e \mathbf{C}^{-1}|}}{\sqrt{2\pi^d}} \exp\left(\frac{1}{2} N_e (\mathbf{M} - \mathbf{M}_e)^t \mathbf{C}^{-1} (\mathbf{M} - \mathbf{M}_e)\right) \\ & \times c(d, N_e) \left(\frac{1}{2} N_e \mathbf{C}_e\right)^{\frac{N_e-1}{2}} |\mathbf{C}^{-1}|^{\frac{1}{2}(N_e-d-2)} \\ & \times \exp\left[-\frac{1}{2} \text{Tr}\left(N_e \mathbf{C}_e \mathbf{C}^{-1}\right)\right] \end{aligned}$$

with:

$$c(d, N_e) = \frac{1}{\pi^{\frac{d(d-1)}{4}} \prod_{i=1}^d \Gamma\left(\frac{1}{2}(N_e-i)\right)}$$

Based on this distribution, we can, for every class i , estimate $p(\mathbf{M}_i, \mathbf{C}_i | \mathbf{M}_{neigh}, \mathbf{C}_{neigh}, N_{neigh})$ where \mathbf{M}_i and \mathbf{C}_i are the i class mean vector and covariance matrix, respectively, and \mathbf{M}_{neigh} , \mathbf{C}_{neigh} and N_{neigh} are, respectively, the empirical mean

vector, empirical covariance matrix and pixel number in the neighborhood or in the segment used.

Classifications based on the Gauss-Wishart (or possibly the Gauss contextual) method provide an average percentage of correctly classified (PCC) pixels in the order of 70 to 80% for a temporal series of at least 3 dates. Raw images can be classified as accurately as filtered images (Kuan MAP filtering; see section 6.3.1.1) with better preserved structures (such filtering always leads to some edge and sharpness loss; see Chapter 6). To yield very good results with filtered images, classification must also be contextual (5×5 neighborhoods, for instance), but this may lead to considerable loss in edge sharpness and fine detail. Between the two of them, the vectorial MMSE filtering and segmentation (see section 6.4.2) can alleviate these problems, especially when dealing with very patchy landscapes, and lead to PCC results of more than 80%.

The PCC results may slightly improve if image segmentation is performed beforehand. Images in this case are visually better and calculation time is reduced more than tenfold. While this principle has a clear advantage over others, it also requires segmentation accuracy that in some regions may be difficult to obtain.

In the case of complex data, where element statistics of a surface scattering matrix are concerned, different distributions have been proposed for measured covariance matrices, depending on whether these matrices are defined for a homogenous or a heterogenous area. When the surface is homogenous, we use the complex Wishart distribution with form parameter L , the number of looks representing the distribution of a complex circular Gaussian vector covariance matrix. The Bayesian classifier proposed in [LEE 94a] uses the Wishart model. Every class is characterized by its mean covariance matrix C_j and form parameter L , which is the same unless classes are at very different angles of incidence and the data are airborne-type with ground coordinate projection (the number of independent samples per pixel varies in this case). This is a generalization of the multi-look single-channel classifier [FRO 85] to the multi-dimensional case. Since classes for all these classifiers are supposed to represent homogenous image surfaces, the only characteristic that makes a difference is the covariance class matrix. In practice, decision rules only depend on local and average values of the backscattering class characteristics.

Let us consider the classification, seen as the ML of a vector y , denoted here by \mathbf{Z} , of complex single-look amplitudes within the same supposedly homogenous area; the dimension of \mathbf{Z} is denoted by d . Its distribution is a complex multi-variate Gaussian distribution. If $\mathbf{C}_i, i \in \{1\dots N\}$, is the complex covariance matrix of class i ,

representing a homogenous surface, the conditional probability density of vector \mathbf{Z} with respect to class i is written (see equation [5.15]):

$$p(\mathbf{Z}|\omega_i) = \frac{1}{\pi^d \det \mathbf{C}_i} \exp \left[-\mathbf{Z}^T \mathbf{C}_i^{-1} \mathbf{Z} \right], \quad [7.1]$$

which makes it possible to define the optimal Bayesian classifier [KON 87].

In the multi-look case, we use distribution of the empirical covariance matrix Σ ; this aspect will be addressed in section 7.4.

7.2.2.2. Unsupervised methods

Parameters can be estimated using the EM and SEM methods the same way as we did for per-pixel methods. The single-channel case with context (neighborhood) was studied in [QUE 97, QUE 93b]. In SAR images, however, statistical estimation based on samples is not always accurate. Perfectly representative class samples are often difficult to achieve. Using pixels with an unknown affiliation class has been proposed as a way to stabilize the statistics of every class [SHA 94]. A number P of pixels of unknown class are added to sample set S of pixels of all classes. The probability of pixels of this set is a law of mixed probability densities of N_ω classes:

$$p(X|\theta) = \sum_1^{N_\omega} a_i p_i(X|\omega_i), \text{ where } \theta \text{ is the set comprised of } a \text{ priori probabilities } a_i \text{ and parameters } C_i \text{ of the } n \text{ conditional class distributions.}$$

Where preliminary segmentation is used in the unsupervised case, the region is assumed to consist of a mix of pixels of different classes. As the EM algorithm is applied, we use pixels from samples of each class together with those in the segment in order to avert that the parameters of class-conditional distributions drift too far away from their initial values. Unless we take this precaution, this algorithm would alter class statistics by adjusting the distributions to those of pixels in the segment. This principle tends to favor classes with the fewest samples. The classes obtained after each alteration do not necessarily correspond to our predefined classes any longer and those for which few learning pixels are available will see their statistics change significantly. This correction pattern only applies if the segment is not dominated by any one predefined class.

7.2.3. Global Bayesian methods

Methods of this type are essentially based on a Markov modeling of class images. Radar imaging is indeed very complex, and the Markov tool is at its best when it shows, as it does here, its capacity of imposing spatial continuity despite a lack of information. The Markov methods were first applied to the classification of noisy images by Geman and Geman [GEM 84]. They were adapted and applied to radar imaging for different types of data (single- and multi-look amplitude images [DER 87, KEL 88, MAR 96], complex images [DER 90, RIG 91] and polarimetric images [RIG 92, SMI 97, LEH 96]).

The Markov methods consist of imposing some degree of regularity to the desired solution x by imposing on X to be a Markov field conditionally to the type of neighborhood selected. The Markov hypothesis makes it possible to bring back the global constraint on $p(X)$ to local constraints governing interactions between labeled pixels.

7.2.3.1. Supervised methods

The parameters, especially those of *a priori* $p(X)$ model, are in most cases either learned beforehand or empirically set.

Selected neighborhoods are often those of 4 and 8 connectedness and the Markov model is a Potts model making it possible to favor the configurations in which labeled pixels belong to the same class. The conditional local distribution is then written:

$$p(x_s = \omega_i / x_r, r \in N_s) = \frac{1}{\zeta} \exp \left[-\beta \cdot \sum_{r \in N_s} \delta_k(\omega_l, x_r) \right] \quad [7.2]$$

where N_s is the considered neighborhood around pixel s , β is a positive constant, δ_k is the Kronecker symbol and ζ is the local normalizing constant called partition function. Classifications provided by this type of model are made up of relatively compact areas, but this comes at a cost for finer structures.

The search for an optimum MAP can be carried out either in a deterministic fashion using iterated conditional modes (ICM) [BES 86] and taking the risk of converging toward a local optimum, or in a stochastic manner by using a simulated annealing technique [GEM 84] to attain global optimum.

The Markov fields are useful due to their high flexibility to accommodate all sorts of constraints (provided that they can be locally expressed). We can introduce, for example, a time dimension in the classification [SHI 96], or use a board process to control the shape of regions with a boundary map defined on the dual grid of that of the pixel [GEM 84] (which may be provided by, for example, a geographic information system [SHI 96]).

Generally speaking, this type of method has the drawback of being slow (although ICM helps step up the search for solutions). Parameter adjustment is also a problem: the more complex the model, the harder it is to adjust regularization parameters in relation to one another.

7.2.3.2. *Unsupervised methods*

In the Gibbs SEM algorithm [CHA 89, MAR 96, DEL 97a], distribution parameters are unknown and have to be estimated on the raw image. Estimating then becomes a more complex problem since we have to find not only image x that optimizes the *a posteriori* probability, but also the law parameters. This algorithm implements a stochastic training and a global model. We choose to estimate *a priori* probabilities of a label for the next iteration based on the percentage of labels drawn from this image during the current iteration. The average of a class is estimated by averaging over observations on pixels to which the label has been assigned.

The principle consists of developing the likelihood logarithm by finding *a priori* and *a posteriori* probabilities in every pixel. Starting from a maximization of log-likelihood under constraint of *a priori* probabilities as being equal to 1, we obtain two equations. Based on these two equations, we will construct an iterative scheme that will allow us to estimate the distribution parameters.

7.3. Application of the Bayesian methods to ERS-1 time series

In temperate regions, optical data are generally well suited to agricultural land mapping. Nevertheless, the accuracy of classification-based maps heavily depends on the acquired image data. The phenologic stages of crops are more easily distinguished when temporal series are available. Since the capacity of taking optical images at precise dates related to crop growth depends on weather conditions, monitoring agencies in charge of these crops are increasingly interested in the opportunities provided by spatial radar sensors.

For illustration purposes, we chose a site near the town of Bourges, France, which is essentially made up of agricultural land plots, woods and forests. The classes we considered were: Forest, Orchard, Wheat, Bread Wheat, Sunflower,

Corn, Barley, Rape, Peas, Clover, Meadow, Ray-grass, Fallow, Bare Soil, Urban Land, Roads and Water.

The test zone was 36 x 14 sq km in area at ground level, with a pixel corresponding to 12.5 x 13 sq m (2,900 line images of 1,100 pixels). The data were as follows:

- 6 ERS-1 single-look complex (SLC) images taken on: March 26, April 30, June 4, July 9, August 13 and September 17, 1993. These data were brought down to multi-look format by 4-pixel spatial averaging in azimuth and by sub-sampling of one in three pixels. The pixel then corresponds to 12.5 m in range and 13 m in azimuth. The spatial multi-look was achieved by means of a spatial whitening filter (see section 6.2.2.1). As a result, speckle statistics precisely corresponded to a four-look generalized Gamma distribution;
- 2 SPOT data of June 9 and August 26, 1993;
- 2 sets of ground-truth data: one gathered in July and another in late September.

ERS-1 data fall into 2 categories:

- raw (unfiltered) amplitude images; and
- vectorial MMSE filtered images: a multi-temporal linear filter estimating in terms of mean square error minimization, coupled with segmentation, as proposed in Chapter 6 (see Figure 7.2b).

In order to improve the adjustment to Gaussian distributions, the data were used in logarithm form. Classifications were performed using maximum likelihood and the Gauss or Gauss-Wishart law for contextual information. The role of this example is to highlight multi-temporal, texture and context contribution. Texture is introduced by using derived channels (local calculation of the standard deviation and coefficient of variation) and context by statistical calculations on pixel neighborhood (sliding window) or segmentation areas.



Figure 7.1. ERS-1 (© ESA) image of the Bourges region resulting from multi-temporal simple pixel-by-pixel summation of six 4-look images obtained through a spatial whitening filter (4×1 windows in azimuth) applied to SLC images

7.3.1. Pixel classification

7.3.1.1. The single-channel case

The average rate of PCC is in the order of 14-17% for raw images and 20-30% for filtered images. These are very low rates, but normal, given our single frequency, single polarization approach². The highest rates correspond to crop maturity periods.

7.3.1.2. Multi-temporal contribution

Classifications have been carried out by adding channels chronologically to assess the importance of multi-temporal data. The study showed an approximate 15% rise in PCC average once the first date had been added. Then the growth percentage of correct classification decreased as yet more dates were added.

² Comparatively, SPOT data have PCC rates of 25-37% when a single spectral band is used.

7.3.1.3. *Introduction of derived channels*

In order to partly avoid class mix-ups, contextual information is introduced through two textural derived channels, one corresponding to the standard deviation and another to the variation factor. These statistics are calculated for every individual date using a sliding window of 7×7 pixels, which provides a good trade-off between estimator roughness and spatial accuracy of the measure for amplitude multi-look ERS images. The contribution of these channels is important in the case of raw single-channel images for which gains of 4-10% are reported. Contextual information is indeed implicitly introduced in these new textural channels. For single-channel filtered images, the gain drops to between 1% and 5%. The more highly filtered the image is, the less important the contribution of these channels becomes.

Improvement due to these derived channels is also found in the multi-temporal case where they enhance date by date classification. The rate of pixels correctly classified is increased by 6.5%, 10%, 11% and 9%, respectively, in raw images for which 1, 2, 3 and 4 dates are available. For highly filtered images, the gain becomes as negligible as 1.5-5% where filtering is carried out with an association of MAP Gamma and Kuan filters. These increases are due to the various phenologic culture stages. The textural standard deviation and coefficient of variation channels provide complementary information. Corn, for example, with a growth that extends over a very long period (from sowing in late March – early April to harvesting in late October – mid-November) and with variations reported throughout these months, has the highest gain in pixels correctly classified.

Coefficient of variations are not really interesting until the 3rd date onwards. On the first dates (late March and late April), the fields are still in bare soil state and textural information has only weakly discriminatory power.

Structures are indeed enhanced and wire-like elements such as roads are more apparent, nevertheless specific algorithms are more appropriate for this type of object (see Chapter 8 and [TUP 99]).

7.3.1.4. *Contribution of contextual classification methods*

Contextual methods make a very important contribution. In single-channel classification, the average percentage of correctly classified (PCC) pixels is now in the order of 25-35% for raw images, which corresponds to a 15% average gain, and of 27-37% for filtered images, or a gain of 2%. Where 6 different dates are used, PCC rates are similar to those obtained for filtered images (around 75%), except for images filtered by vectorial MMSE (associated with segmentation), where the PCC rate is close to 85%. These results clearly indicate that pixel classification is ill adapted to raw data because of speckle and that contextual classifications are equally

advisable for filtered data. The sliding window for this type of (4-look) image is optimally sized 7 x 7 pixels for raw data and 5 x 5 pixels for filtered data (see Table 7.1).

Number of dates used	Classification on unfiltered image	Classification on vectorial MMSE filtered image after segmentation
1 date	27.7%	
2 dates	40.7%	49.9%
3 dates	54.3%	65.7%
4 dates	64%	74.3%
5 dates	69%	77.4%
6 dates	74%	82.2%

Table 7.1. Performances of multi-temporal contextual classification, as evaluated by the rate of pixels correctly classified: PCC. For the filtered image, multi-temporal alone is calculated

The classifications map based on vectorial MMSE filtered images (see Figure 7.3 (top)) shows some confusion on the edges. Pixels outlining the forest are misclassified as corn or other classes, which does not happen when classification relies on raw data and also introduces segmentation, where the context of a pixel is a region integrating homogenous areas (see Figure 7.3 (bottom)). However, too many pixels are classified as forest and the PCC rate is slightly lower than that of the previous classification. Some small details are not present. The problem is to correctly adjust the size of segmentation during labeling. A fine segmentation is advisable for classification purposes. If segmentation is too large, the speckle effect is indeed suppressed, but small details and structures are lost.

The contribution from derived channels (standard deviation and coefficient of variation) is 5% lower in the contextual case. This may be due to how these textured channels are calculated, i.e., by integrating the context through a sliding window. As a result, the information that we gather is implicitly despeckled. This leads to an enhanced PCC rate in raw images, particularly where pixel methods are used. By contrast, the gain is slim in images already filtered. Adding these channels mainly makes it possible to improve the structures.

The classification rates obtained from an ERS-1 time series by using context in both cases cited above (vectorial MMSE filtered data and segmentation support) are close to those that might be obtained from optical data.

Several supervised classification methods have been tested; among them, those that relied on the Gauss or Gauss-Wishart laws yielded the best results. Images obtained with EM and Gamma texture methods needed filtering, but the results were often too noisy. The MAP method was used for classification and yielded satisfactory results, though a large number of iterations (around 40) were necessary due to the very slow temperature variation of simulated annealing and segmentation was needed for edge constraint.

Examples of classification results from a 6-date multi-temporal series are shown in Figure 7.3. The ERS-1 image of the Bourges region, as obtained by mere pixel-by-pixel addition of 4-look images acquired on 6 different dates is shown in Figure 7.1. Figure 7.2 provides intermediate results which can be used to classify only part of Figure 7.1 (corresponding to the lower right area). The edge image (see Figure 7.2a) corresponds to a radiometric segmentation obtained by averaging over the 6 dates with edge detector ROEWA (or “ratio of exponentially weighted averages”; see section 8.2.3), followed by calculating the watershed line (see [SCH 94]) and the fusion of regions. The edges are eventually repositioned with the aid of Markov fields. Figure 7.2b shows the linear vector MMSE estimator, where statistical parameters are calculated over entire regions produced by radiometric segmentation with the method described above. Proper classification results are obtained by multi-temporal classification to the maximum likelihood by a Gauss-Wishart law: on regions coming from radiometric segmentation on the one hand (see results at the bottom of Figure 7.3) and, on the other hand, by a contextual method applied to data yielded by the linear vector MMSE estimator (results shown at the top of Figure 7.3).

7.3.2. Radar and SPOT fusion, comparison with SPOT

ERS data often have a poor capacity for discriminating crops. Better results are generally obtained by a simple non-contextual classification applied to optical (e.g., SPOT) images on particular dates carefully selected according to the phenologic crop cycle. However, for weather reasons, we do not always have this option.

ML classifications combining both types of data, SPOT and ERS-1, in amplitude logarithm form lead to better results than each of them would yield alone. These classifications use data fusion techniques that make it possible to combine a Gaussian ML method for SPOT data with a contextual method for ERS-1 data. The average rate of the correct PCC classification for SPOT is in the order of 75-85%. A fusion with raw radar images using the Gauss-Wishart method improves SPOT results by about 5-14% on average, depending on conditions. Structures in the land plot image are improved.

7.3.3. Conclusion

The results given above were obtained from an example. Many other scenes have been tested, however, including agricultural plots, temperate forests and Mediterranean areas, as well as data from different sensors (RADARSAT, JERS-1), and the results were in the same order of magnitude.

Pixel classification shows the necessity of preliminary data filtering. Textural derived channels based on the standard deviation and coefficient of variation increase the degree of crop discrimination, particularly those like corn, where the vegetation cycle is covered by the timescale under investigation. The gain is higher for raw images than it is for filtered images and the pixel-based case.

Raw SAR data classified by contextual methods and possibly textural derived channels can lead to similar results as are obtained using traditional MAP or Kuan filtered data, or by multi-temporal addition combined with, for example, a Gamma-MAP spatial filter, and structures are retained well by these data. The best results, however, are obtained from vectorial MMSE filtered data associated with segmentation, or from raw data by introducing segmentation in the classification process. This is an extremely fast tool for obtaining satisfactory results.

The SPOT/SAR fusion process makes it possible to remove to a certain extent confusions generated by SPOT alone by improving structure preservation and ensuring a gain of about 5-14%, depending on image type and classes involved.

Significant research projects are still underway in these areas. They are particularly aimed at achieving classifications directly from complex data which, unlike amplitude data, include full information and which we know how to model reasonably well (circular Gaussian model).

7.4. Classification of polarimetric images

As images with vector data in every pixel, polarimetric images are *a priori* better suited for classification than scalar images. Methods to classify polarimetric SAR images fall into two broad types:

- methods based on an analysis specific to polarimetric information. They mostly rely on identifying the backscattering mechanisms by means of specific discriminators [TOU 92] and on interpreting these mechanisms in terms of backscattering processes. The best known methods of this kind are those proposed by Van Zyl [VAN 89] and by Cloude and Pottier [CLO 96]; and

– methods applying “traditional” classification criteria to the distribution of polarimetric SAR data ([RIG 92], etc.).

The main advantage of the former is that they are general and do not require training data: although entirely supervised, they are automatic methods, as the researched classes are known *a priori*. Considering the latter, their advantage lies in being more refined and able to distinguish among a higher number of classes.

7.4.1. Classification in three major backscattering mechanisms [VAN 89]

7.4.1.1. The principle

Van Zyl identifies three major backscattering mechanisms illustrated in Figure 7.4:

- a backscattering that involves an odd number of reflections;
- a backscattering that involves an even number of reflections; and
- a diffuse backscattering.

The advantage of considering these backscattering properties is that they very easily translate into polarization properties of the backscattered wave. Indeed, according to [VAN 89], in the case of a single reflection, S_{hh} and S_{vv} are in phase. Each new reflection then adds 180° to the phase difference between S_{hh} and S_{vv} . The classification is therefore mainly based on the phase difference between the co-polar channels *hh* and *vv*. If we denote by $\arg(S_{hh} \cdot S_{vv}^*)$ the phase difference between *hh* and *vv*, we then have:

- for class 1 (odd number of reflections): $\arg(S_{hh} \cdot S_{vv}^*) \approx 0^\circ$;
- for class 2 (even number of reflections): $\arg(S_{hh} \cdot S_{vv}^*) \approx 180^\circ$; and
- for class 3 (diffuse backscattering): S_{hh} and S_{vv} are decorrelated and have any phase difference.

7.4.1.2. Implementation and illustration of results

Under the double assumption of reciprocity and decorrelation of co- and cross-polar channels, the physical properties of the different backscattering types considered by Van Zyl translate into the following relations involving the covariance terms:

$$\begin{aligned}
 \arg(S_{hh}S_{vv}^*) \approx 0^\circ &\Rightarrow \langle \operatorname{Re}(S_{hh}S_{vv}^*) \rangle > \varepsilon \\
 \arg(S_{hh}S_{vv}^*) \approx 180^\circ &\Rightarrow \langle \operatorname{Re}(S_{hh}S_{vv}^*) \rangle < -\varepsilon \\
 \arg(S_{hh}S_{vv}^*) \text{ any} &\Rightarrow \langle \operatorname{Re}(S_{hh}S_{vv}^*) \rangle \approx 0 \Leftrightarrow \lvert \langle \operatorname{Re}(S_{hh}S_{vv}^*) \rangle \rvert < \varepsilon
 \end{aligned}$$

where $\langle z \rangle$ is the average z over a set of spatially “neighboring” pixels and ε is a positive constant *a priori* small. In fact, [VAN 89] took $\varepsilon = |S_{hv}|^2$, thus authorizing relatively large phase deviations. He actually had to do so to make his classification applicable, since due to noise, speckle in particular, phases coming from the same homogenous area (distributed target) may see their distributions differ by up to 50% in standard deviation.

In [VAN 89], the author also proposed another version of his classification algorithm, this time relying on backscattered wave properties, particularly the polarization variations of the backscattered wave versus the transmitted wave properties considering either linear or circular polarization. The advantage of this second implementation lies in the fact that it reveals some pixels that are not classified.

Finally, since the method assumes that co- and cross-polar channels are decorrelated (see section 5.5.4), a variant of this method has therefore been proposed [MAS 94] to preliminarily check this assumption and take into account, if necessary, a new class of backscattering mechanisms: that of targets with azimuthal asymmetry.

Figures 7.5a and 7.5b illustrate the results of the classification by Van Zyl at L-band and C-band respectively, of a French agricultural site at Orgeval that also included a forest area and a few very small urban areas (hamlets). The L-band yielded the best results; it is in fact the band used by Van Zyl to present his results (1989). The towns (classified as “even number of reflections”) are accurately recognized, the forests are classified as “diffuse backscattering” and the fields are assigned to the “odd number of reflections” class (or “diffuse backscattering” for some). At C-band the results are not so good: the towns are less well recognized and forests less well delimited. Forests are in the “odd number of reflections” class, while the fields are mostly classified as “diffuse backscattering”. This difference in terms of backscattering mechanism between L (wavelength ≈ 24 cm) and C (wavelength ≈ 6 cm) bands is due to the different proportion between scatterer size and wavelength and different penetration depth through vegetation.

We conclude that this first classification method is, on the one hand, very simple to implement and entirely automatic and, on the other hand, provides a simple physical interpretation of the interaction between radar wave and environment. It is nevertheless a rough method, since two different environments may have the same backscattering mechanisms (out of the three defined by Van Zyl) and since each class includes areas quite different from one another in terms of physical characteristics.

7.4.1.3. Conclusion

To conclude, we will refer to the three-component model of Freeman and Durden [FRE 92a], which tackles the three major backscattering mechanisms identified by Van Zyl. Freeman and Durden write the observed covariance matrix as a weighted sum of three matrices, every one of which is associated with one of the following mechanisms: (i) bare soil reflection; (ii) dihedral reflection and (iii) volume scattering (a cluster of randomly oriented dipoles). Van Zyl's "diffuse scattering" class then corresponds to the case in which none of these mechanisms predominates (while either mechanism (i) or mechanism (ii) is prevalent in the two other Van Zyl classes). Therefore, the main merit of [FRE 92a] consists of making a distinction between the cases in which a backscattering mechanism predominates (and then we can *a priori* identify as many classes as there are mechanisms) and the case where neither mechanism is prevalent (and where we therefore have a sole class we will call "diffuse"). This is the very principle of the entropy-based classification that is discussed below.

7.4.2. Entropy-based classification [CLO 96]

7.4.2.1. The principle

The classification method developed in [CLO 96] is based on estimating the backscattering complexity degree and, if possible, identifying the predominant backscattering mechanism. Instead of writing the observed coherence matrix \mathbf{T} (see section 7.4.2.2, equation [7.3]) as the sum of a matrix corresponding to an identified backscattering mechanism and a noise matrix (as in the Huynen decomposition), the authors suggest analyzing \mathbf{T} in terms of vectors and eigenvalues, in order to identify the part (in terms of contribution to the observed response) of every backscattering mechanism involved.

The method therefore consists of writing the coherence matrix \mathbf{T} in the form of a product $\mathbf{U} \cdot \Delta \cdot \mathbf{U}^{-1}$, where Δ is the diagonal matrix of researched eigenvalues and \mathbf{U} the matrix of eigenvectors. The backscattering complexity degree will then be represented by the entropy associated with the eigenvalues and the predominant backscattering mechanism by either the first eigenvector or the weighted average of

eigenvectors [CLO 96]. Besides, note that in this classification, the azimuthal symmetry assumption is no longer necessary.

7.4.2.2. Implementation and illustration of results

The vectorial writing of complex backscattering matrix \mathbf{S} actually corresponds to its decomposition on the basis $\begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ \sqrt{2} & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & \sqrt{2} \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & \sqrt{2} \end{pmatrix}$, whereas the Huynen decomposition theorems use a different basis:

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

The corresponding vector \mathbf{S} is then:

$$\overrightarrow{\mathbf{k}_P} = \frac{1}{2} \cdot (S_{hh} + S_{vv} \quad S_{hh} - S_{vv} \quad S_{hv} + S_{vh} \quad i(S_{hv} - S_{vh}))^t,$$

which under the reciprocity assumption reduces to:

$$\overrightarrow{\mathbf{k}_P} = \frac{1}{2} \cdot (S_{hh} + S_{vv} \quad S_{hh} - S_{vv} \quad 2 \cdot S_{hv})^t$$

and coherence matrix \mathbf{T} is written:

$$\mathbf{T} = \overrightarrow{\mathbf{k}_P} \cdot \overrightarrow{\mathbf{k}_P}^{*t} = 1/4. \quad [7.3]$$

$$\begin{pmatrix} |S_{hh}|^2 + |S_{vv}|^2 + 2\text{Re}(S_{hh}S_{vv}^*) & |S_{hh}|^2 - |S_{vv}|^2 - 2\text{Im}(S_{hh}S_{vv}^*) & 2(S_{hh}S_{hv}^* + S_{vv}S_{hv}^*) \\ |S_{hh}|^2 - |S_{vv}|^2 + 2\text{Im}(S_{hh}S_{vv}^*) & |S_{hh}|^2 + |S_{vv}|^2 - 2\text{Re}(S_{hh}S_{vv}^*) & 2(S_{hh}S_{hv}^* - S_{vv}S_{hv}^*) \\ 2(S_{hv}S_{hh}^* + S_{hv}S_{vv}^*) & 2(S_{hv}S_{hh}^* - S_{hv}S_{vv}^*) & 4|S_{hv}|^2 \end{pmatrix}$$

\mathbf{T} can also be easily obtained starting from either Mueller matrix \mathbf{M} (see equation [1.22]) or covariance matrix \mathbf{C} .

The decomposition of \mathbf{T} revealing eigenvectors and eigenvalues is written:

$$\mathbf{T} = \mathbf{U}_3 \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \cdot \mathbf{U}_3^{*t} \quad [7.4]$$

where \mathbf{U}_3 is the unitary matrix of dimension 3 of the eigenvectors:

$$\mathbf{U}_3 = \begin{bmatrix} \cos \alpha_1 & \cos \alpha_2 & \cos \alpha_3 \\ \sin \alpha_1 \cdot \cos \beta_1 \cdot e^{i\delta_1} & \sin \alpha_2 \cdot \cos \beta_2 \cdot e^{i\delta_2} & \sin \alpha_3 \cdot \cos \beta_3 \cdot e^{i\delta_3} \\ \sin \alpha_1 \cdot \cos \beta_1 \cdot e^{i\gamma_1} & \sin \alpha_2 \cdot \cos \beta_2 \cdot e^{i\gamma_2} & \sin \alpha_3 \cdot \cos \beta_3 \cdot e^{i\gamma_3} \end{bmatrix}$$

NOTE: generally, not all parameters $(\alpha_1, \beta_1, \delta_1, \gamma_1, \alpha_2, \beta_2, \delta_2, \gamma_2, \alpha_3, \beta_3, \delta_3, \gamma_3)$ are independent; in particular, by assuming the decorrelation of the co- and cross-polar channels, the number of independent parameters is reduced to 3.

Once the eigenvalues and eigenvectors of \mathbf{T} have been obtained, Cloude and Pottier propose a classification based on the two following parameters [CLO 96]:

– entropy H associated with eigenvalues:

$$H = \sum_{i=1}^3 -P_i \cdot \log_3(P_i) \text{ with } P_i = \lambda_i / \left(\sum_{i=1}^3 \lambda_i \right) \quad [7.5]$$

– parameter α , which is equal to α_1 if a predominant backscattering mechanism can be identified, or is estimated as the weighted average of the mechanisms involved: $\alpha = \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3$.

The main backscattering mechanisms that can then be identified are:

- surface backscattering;
- dipole backscattering; and
- multiple backscattering (including, for example, dihedral backscattering).

When H is low, it is simple scatterers, or all scatterers oriented in the same direction that are responsible for these mechanisms. By contrast, an increase in H reflects a higher number of scatterers with different properties. Besides, according to the definition of α , the lower H is, the more relevant the potential of α for

identifying the predominant backscattering mechanism and hence the more numerous the mechanisms that can be distinguished. Conversely, the higher H is, the fewer backscattering types that can be discerned. Finally, if H is very high, some of these mechanisms cannot be distinguished at all.

In [CLO 96], for example, 9 classes are proposed: 3 classes distinguished by α combined with 3 classes distinguished by H , except that one of these, i.e. surface backscattering with high entropy, cannot be observed.

Examples of the Cloude and Pottier classification results at L-band and C-band are shown in Figures 7.6a and 7.6b, respectively. The site is Orgeval again, essentially agricultural and forested areas including few hamlets. As in Van Zyl's classification, we notice the L- and C-bands differ in terms of the backscattering mechanisms involved. At L-band, for example, most of the forest belongs to the class of "high-entropy volume backscattering" (dense vegetation), while a small part of it (especially trunks and branches) falls into the "high-entropy multiple backscattering" class. At C-band, it is essentially divided into two classes with average entropy: "dipole backscattering" and "surface backscattering". These results are easily explicable considering the difference of wave penetration depth in the land cover (which is in the order of several wavelengths).

7.4.3. Polarimetric Bayesian classifications

7.4.3.1. The single-look Gaussian case

Rignot and Chellappa proposed applying a Bayesian classification method, namely the MAP criterion, to classify polarimetric SAR images [RIG 92]. They assumed a conditional multi-variate Gaussian distribution of $\tilde{\mathbf{S}}$ (equation [5.15]), which under the double assumption of reciprocity ($S_{hv} = S_{vh}$) and decorrelation of the co- and cross-polar channels $(\langle S_{hh} \cdot S_{hv}^* \rangle = \langle S_{hv} \cdot S_{vv}^* \rangle) \approx 0$ is written:

$$p(\tilde{\mathbf{S}} / x_s = \omega_l) = \frac{\exp \left\{ -\frac{|S_{hh}|^2}{\tilde{\sigma}_{hh,l} \cdot (1 - |\tilde{\rho}_l|^2)} - \frac{|S_{hv}|^2}{\tilde{\sigma}_{hv,l}} + A \right\}}{\pi^3 \cdot \left[\tilde{\sigma}_{hh,l} \cdot \tilde{\sigma}_{hv,l} \cdot \tilde{\sigma}_{vv,l} \cdot (1 - |\tilde{\rho}_l|^2) \right]} \quad [7.6]$$

$$A = -\frac{|S_{hh}|^2}{\tilde{\sigma}_{vv,l} \cdot (1 - |\tilde{\rho}_l|^2)} + -\frac{2 \cdot \operatorname{Re}(S_{hh} \cdot S_{vv}^* \cdot \tilde{\rho}_l^*)}{\sqrt{\tilde{\sigma}_{hh,l} \cdot \tilde{\sigma}_{vv,l}} \cdot (1 - |\tilde{\rho}_l|^2)}$$

$$A = -\frac{|S_{vv}|^2}{\tilde{\sigma}_{vv,l} \cdot (1 - |\tilde{\rho}_l|^2)} + \frac{2 \cdot \operatorname{Re}(S_{hh} \cdot S_{vv}^* \cdot \tilde{\rho}_l^*)}{\sqrt{\tilde{\sigma}_{hh,l} \cdot \tilde{\sigma}_{vv,l}} \cdot (1 - |\tilde{\rho}_l|^2)}$$

where $\tilde{\sigma}_{pq,l} = \left\langle |S_{pq}|^2 \right\rangle_{l-class}$, with pq taking the value hv or vv for hh , and $\tilde{\rho}_l = \frac{\left\langle S_{hh} \cdot S_{vv}^* \right\rangle_{l-class}}{\sqrt{\tilde{\sigma}_{hh,l} \cdot \tilde{\sigma}_{vv,l}}}$ is the correlation coefficient of co-polar channels.

The interactions between labeled pixels are modeled in terms of Markov fields by a simple Potts model (equation [7.2]).

Also in [RIG 92], an unsupervised version of the classification was proposed; it consists of a preliminary class feature estimation by the fuzzy c-means method. The considered feature vector \vec{X} then includes (non-zero) terms of the covariance matrix, i.e., terms $\tilde{\sigma}_{hh,l}$, $\tilde{\sigma}_{hv,l}$, $\tilde{\sigma}_{vv,l}$ and $\tilde{\rho}_l$ involved in equation [7.6]. It was proposed:

$$\vec{X} = \begin{pmatrix} 10 \cdot \log_{10} \left\langle |S_{hh}|^2 \right\rangle \\ 10 \cdot \log_{10} \left\langle |S_{hv}|^2 \right\rangle \\ 10 \cdot \log_{10} \left\langle |S_{vv}|^2 \right\rangle \\ 10 \cdot \log_{10} \left| \left\langle S_{hh} \cdot S_{vv}^* \right\rangle \right| \\ \frac{10}{\ln(10)} \cdot \arg \left\langle S_{hh} \cdot S_{vv}^* \right\rangle \end{pmatrix} \quad [7.7]$$

The main reason for switching to the logarithmic domain was to make the speckle additive (rather than multiplicative as we have in the linear domain). Factor $10/\ln(10)$ preceding the argument is aimed at ensuring similar dynamics on all components of $\vec{\mathbf{X}}$.

Examples of the results of this unsupervised classification at L- and C-bands are shown in Figures 7.7a and 7.7b, respectively. Since the method is unsupervised and based on purely mathematical criteria, we cannot deduce any interpretation straight from the results. We can, however, underline the visible quality of these results, as fields are clearly represented. Concerning crop discrimination and identification (obtained *a posteriori* from test areas, for example), it is the L-band that provides the best results (as a matter of fact, it is precisely at this band that the results are given in [RIG 92]). These results are quite satisfactory, since the prevailing land cover types (forest and main crops) are perfectly well identified, and ambiguity only persists for some minor types of cover.

7.4.3.2. The multi-look Gaussian case

In the multi-look case, [RIG 92] assumes a decorrelation of looks (the authors themselves admit this is quite certainly a false assumption in most cases, particularly when the multi-look image is obtained by plain spatial averaging, but they argue has only little impact on the final results). Based on this assumed decorrelation, they find an expression similar to equation [7.6] (except for factor $1/\pi^{3(L-1)}$ where L is the number of looks).

In a more rigorous approach, it is the Wishart distribution (equation [5.17]) that we have to use, as it is used in [LEE 94b], for example. Expression [7.6] is then replaced by the following, more general expression (equation [5.8]):

$$p(\Sigma / x_s = \omega_l) = \frac{|\Sigma|^{L-d} \exp\{-L\text{tr}(\mathbf{C}_l^{-1}\Sigma)\} K(L, d)}{|\mathbf{C}_l|^L} \quad [7.8]$$

where \mathbf{C}_l is the covariance matrix associated with class l ; Σ is the observed L -look covariance matrix; and $K(L, d)$ is a constant:

$$K(L, d) = \frac{L^L}{\pi^{d(d-1)/2} \Gamma(L) \dots \Gamma(L-d+1)}$$

7.4.3.3. The product model case

Finally, note that whereas the Gaussian hypothesis is by far the most commonly used, other Bayesian classification methods based on K distribution (non-homogenous texture distributed according to a Gamma law) have been tested [NEZ 96, SER 97]. These are more general and sophisticated methods, but they are also more delicate to implement. Above all, estimating form parameters of the Gamma laws may lead to errors that would wipe out any advantage of using the K law instead of the Gaussian approximation. A proposition has been made in [SER 97] to use the Wishart distribution conditioned to texture μ :

$$p(\boldsymbol{\Sigma} / \xi = l) = \frac{|\boldsymbol{\Sigma}|^{L-d} \cdot \exp\left\{-L \cdot \frac{1}{\mu} \text{tr}\left(\mathbf{C}_l^{-1} \cdot \boldsymbol{\Sigma}\right)\right\} K(L, d)}{\mu^{Ld} \cdot |\mathbf{C}_l|^L} \quad [7.9]$$

where the notations are the same as in equation [7.8].

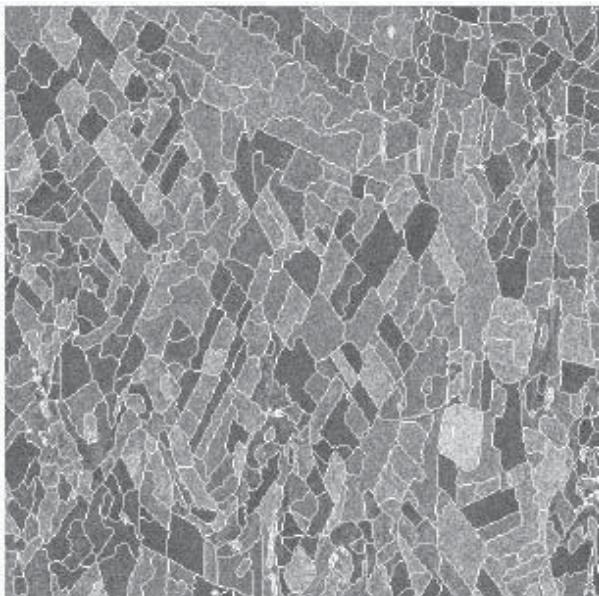
Therefore, Séry in his approach assumed the texture to be known in every pixel and set it aside, settling instead for the classification in [LEE 94a].

Other approaches used texture parameter information to determine the class of a pixel. Examples of such methods are presented in the following section.

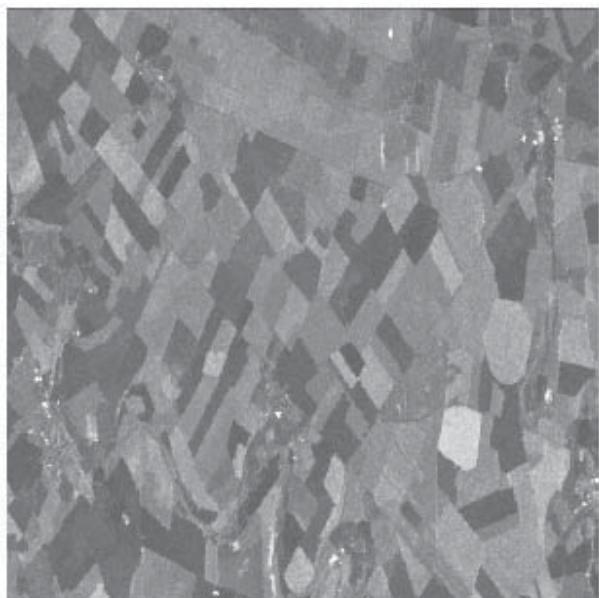
7.4.4. Polarimetric non-Bayesian classifications

Here, we only mention that there are many other classification algorithms aside from the above classification methods which are based on either physical or mathematical signal analysis. These other methods are largely empirical and their performance is often based on rules. We can thus cite the hierarchical classifier proposed by [PIE 94] and another one used by [FRE 94]. Such classifiers generally use multi-band data and rely not only on the different covariance matrix terms (at every band), as in the Bayesian case, but also on features estimated indirectly from polarimetric terms, such as notably the coherence degree of the various channels [TOU 92], or texture.

Finally, from among the methods based on learned rules, let us mention those that rely on neural networks, e.g., [HAR 94] or [CHE 96].



(a) Original image (colored composition) and segmentation



(b) Filtered image

Figure 7.2. Colored composition of images from 3 dates obtained by the linear vector MMSE estimator (b) where the statistical parameters are calculated over entire regions produced by radiometric segmentation (a) (see color plate section)

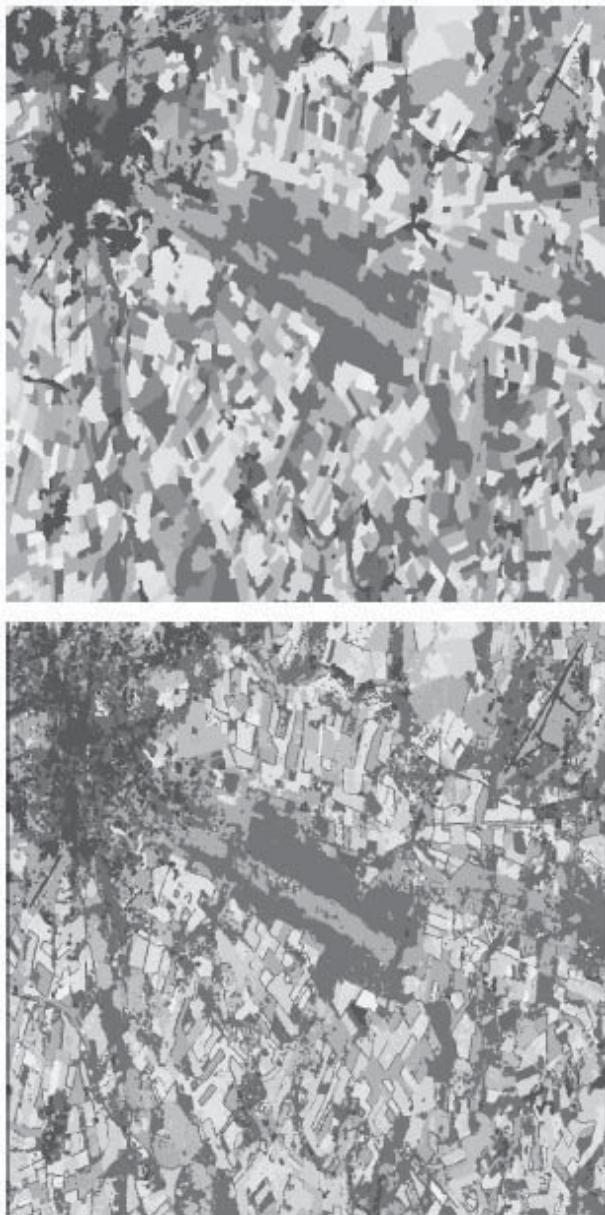


Figure 7.3. Gauss-Wishart multi-temporal ML classification over regions produced by radiometric segmentation (top) and Gauss-Wishart multi-temporal ML contextual (5×5 pixels) classification on data obtained using the linear vector MMSE estimator (bottom) (see color plate section)

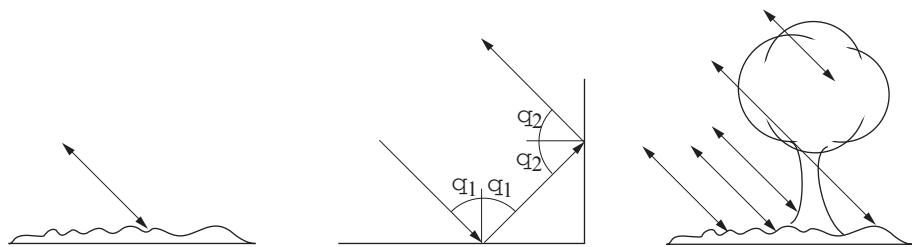
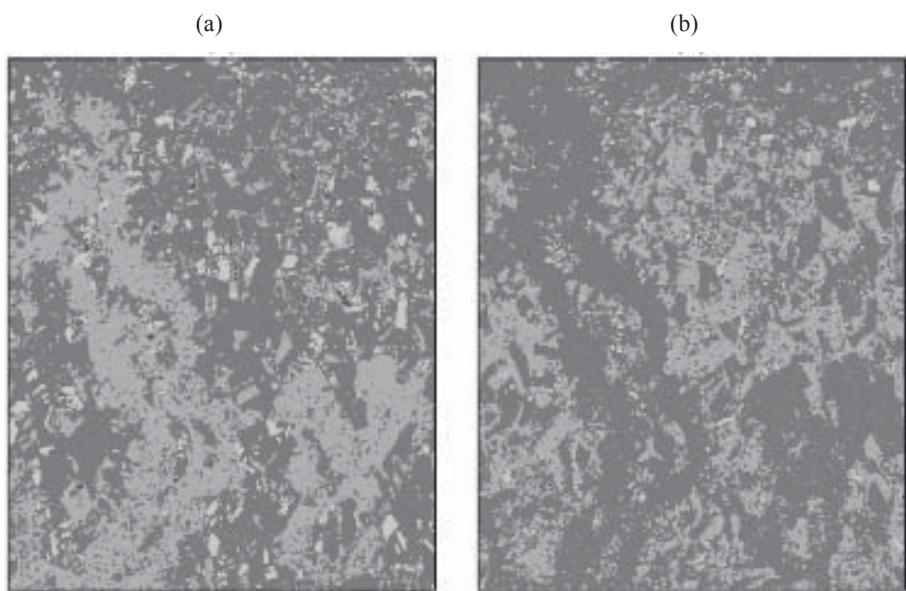


Figure 7.4. Chief backscattering mechanisms according to Van Zyl: (a) odd backscattering; (b) even backscattering; (c) diffuse backscattering



■ odd number of reflections ■ even number of reflections ■ diffuse backscattering □ azimuthal asymmetry

Figure 7.5. Van Zyl classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data (see color plate section)

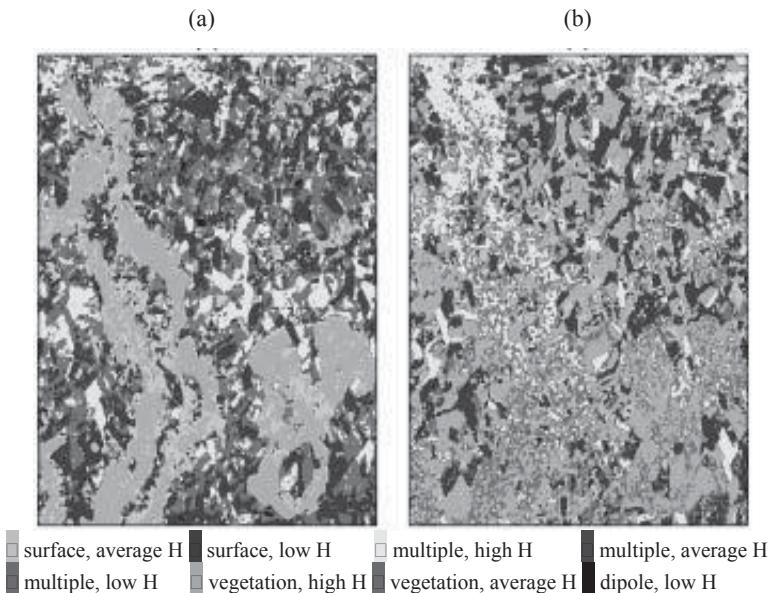


Figure 7.6. Cloude and Pottier classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data (see color plate section)

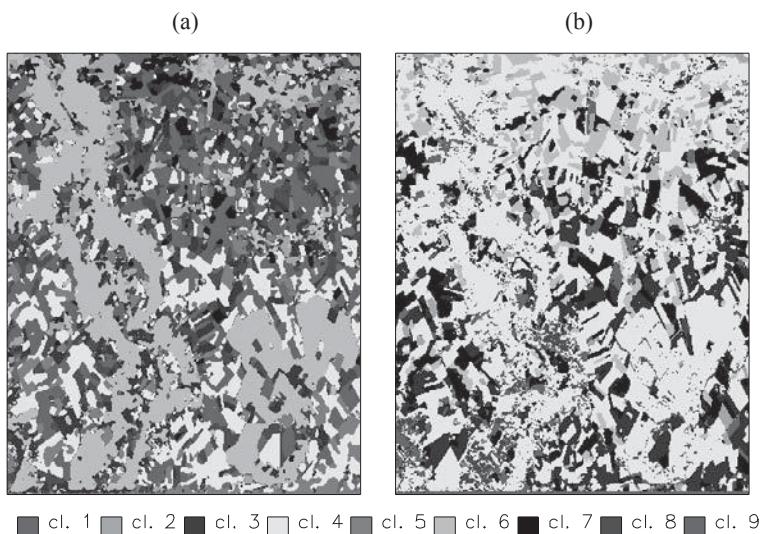


Figure 7.7. Rignot classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data (see color plate section)

This page intentionally left blank

Chapter 8

Detection of Points, Contours and Lines

Numerous remote sensing applications require the detection of elements structuring the image. Among these, shiny targets or points, contours and lines are preeminent. In fact, strong reflectors corresponding to objects are often manmade; lines can be described in numerous themes, whether they are the road and hydrographic networks or geological structures. As for contours, they make dividing the image into areas of interest possible, such as parcel plans, or they characterize regions in relation to their density, such as in urban areas. This chapter describes the techniques making the detection of these three types of primitive elements on radar images possible.

Despite its strong image readability degradation, speckle is interesting because it is statistically well modeled, as we have seen in Chapter 5. This aspect is important and the majority of methods presented in this chapter are based on statistical analysis for the development of detection criteria. Thus, considering a measure m , which indicates the presence, or not, of a target, a contour or a line, we will try to establish the conditional distribution of m according to magnitudes implied: contrast between target and background, or contrast on one side and the other of the contour, or size of the neighborhood used. Then, we will be able to predict, according to the threshold values on m , detector performances in terms of detection probability P_d and false alarm P_{fa} . When it is not possible to obtain an analytical expression of the distribution, simulations may be used in order to obtain P_d and P_{fa} .

8.1. Target detectors

Certainly, target detection has an important place in radar imaging, especially in a military context. The notion of constant false alarm rate detectors (CFAR detectors)¹ is particularly important in radar imaging. In fact, as we have seen in Chapter 5, one of the characteristics of this type of imaging lies in the fact that the standard deviation of a physically homogenous area increases with its average radiometry. Consequently, for inappropriate detectors (such as those adapted to additive noise), the false alarm rate increases in areas with high radiometry. In order to obtain a constant behavior irrespective of local radiometry, it will be necessary to use variable thresholds, which makes the use of these detectors more complex. Instead, as we will see, most of the detectors in radar imaging are based on a ratio.

Targets are characterized by very strong backscattering and they correspond to bright points on the image. Depending on resolution and wavelength, these are typically vehicles, ships, buildings or metallic constructions (bridges, tankers, masts, etc.).

8.1.1. Intensity threshold

If we place ourselves in a Bayesian framework, the problem of target detection is the confrontation of 2 hypotheses: data x belongs either to target C , or to background (texture) F of the image. Ω is the binary random variable representing the class; we then compare probabilities $P(\Omega = C | X = x)$ and $P(\Omega = F | X = x)$. Since:

$$\frac{P(C|x)}{P(F|x)} = \frac{p(x|C)P(C)}{p(x|F)P(F)}$$

and considering that the *a priori* $P(\Omega = C)$ and $P(\Omega = F)$ are generally unknown, we are limited to the criterion of maximum likelihood (ML) by comparing the following ratio² to 1:

¹ A detector is considered CFAR (*Constant False Alarm Rate*) when its false alarm rate is constant on all homogenous areas of the image, i.e. independently of a region's local radiometry.

² It is necessary to consider a interval of affiliation $\Delta(x)$ around x or to discretize data – which is performed naturally in the gray level case – in order to replace probability densities by probabilities.

$$\frac{P(X \in \Delta(x)|C)}{P(X \in \Delta(x)|F)}$$

In the absence of the target's distribution $p(x|C)$ knowledge, we are satisfied with detecting the abnormal points where the probability of appearance is lower than a given threshold. Thus, we will decide that this has to do with a target when $P(X \in \Delta(x)|F) < s$, where s is determined in order to ensure a low false alarm rate. As in the case of strong reflectors,

$$P(X \in \Delta(x)|F) < s - x > t$$

we deduce the false alarm probability:

$$P_{fa} = \int_t^{+\infty} p(x|F) dx$$

A first detection method of targets therefore consists of thresholding pixels amplitudes according to their probability of appearance in the radar image. If the image is homogenous, we can use global thresholding. In this case we consider that the background is defined by the entire image and probability density $p(x|F)$ and threshold are estimated once and for all. Performances then strongly depend on global statistics of the treated image to the detriment of sturdiness.

Thus, we can consider local thresholding, the background being defined in a neighborhood of the pixel for which the threshold is deduced. The neighborhood used may introduce a neutral area, as in Figure 8.1a, in order to ignore the impulse response of the sensor, or in order to model it as in Figure 8.1b. Instead of only comparing the pixel value to the threshold, we can then use the intensity average on the neighborhood, which is defined by a cross-shaped target [LOP 93a]. Probability density $p(x|F)$ is modified when empirical intensity methods for x and F are used and this depends on the number of samples used. It is then necessary to take this into account when calculating the threshold corresponding to a fixed false alarm rate.

This approach is used in [CAV 93, GOU 96] for the detection of urban environment. Let us note that this type of method, and more generally target detection, is very convenient for isolated points but will be less adapted to urban density in the case of satellite resolutions. The choice of global or local thresholding has a great influence on the results obtained, but also on calculation time. In the presence of dark structures, such as water, for example, a local threshold may prove inappropriate (numerous false alarms at area boundary).

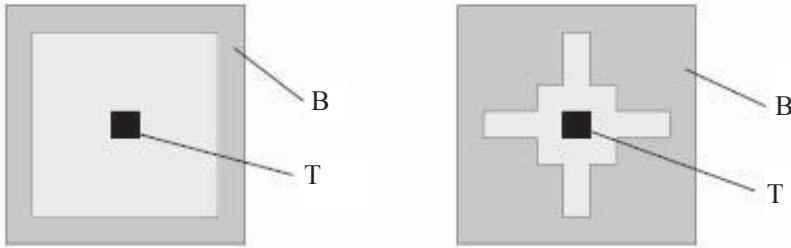


Figure 8.1. Examples of neighborhoods defining target (T) and background (B). The light gray part corresponds to a neutral area and it is not used for statistical calculation

8.1.2. Ratio use

Another approach consists of confronting the two following hypotheses on a pixel neighborhood: either there are 2 areas, target and background, or the neighborhood represents one and the same area. Let us note by V_1 neighborhood 1 (for example, the target), constituted of N_1 pixels and of reflectivity R_1 , by V_2 the neighborhood 2 of N_2 pixels and of reflectivity R_2 and by V the global neighborhood of $N = N_1 + N_2$ pixels and of reflectivity $R = \frac{1}{N_1+N_2}(N_1R_1 + N_2R_2)$. The likelihood ratio is calculated by supposing independent pixels:

$$\log \left[\frac{P(I_i, \dots, \forall i \in V_1 | R_1) P(I_i, \dots, \forall i \in V_2 | R_2)}{P(I_i, \dots, \forall i \in V | R)} \right]$$

which is written, by noting empirical intensity averages \hat{I}_i on V_i and performing the hypothesis of single-look Gamma laws:

$$\begin{aligned}\lambda(N_1, N_2, R_1, R_2, R) = & -N_1 \left(\ln R_1 + \frac{\hat{I}_1}{R_1} \right) - N_2 \left(\ln R_2 + \frac{\hat{I}_2}{R_2} \right) \\ & + (N_1 + N_2) \left(\ln R + \frac{\hat{I}}{R} \right)\end{aligned}\quad [8.1]$$

Note that this model is not verified in the case of targets that do not follow a Gamma law (see equation [5.5]). Similarly, this criterion will be used only as an anomaly test, as will be seen below.

The optimal Neyman-Pearson test given by equation [8.1] requires the knowledge of all parameters R_i and N_i . In practice, R_i values are unknown and the generalized likelihood ratio is used. This is obtained by replacing them with their corresponding ML estimators \hat{I}_i :

$$\hat{\lambda}(N_1, N_2) = -N_1 \ln \hat{I}_1 - N_2 \ln \hat{I}_2 + (N_1 + N_2) \ln \hat{I} \quad [8.2]$$

It is necessary to know distribution $\hat{\lambda}$ in order to be able to calculate the threshold corresponding to a given false alarm probability. In practice, however, it is revealed to be quite difficult, even impossible, to express it analytically. A solution is that of writing $\hat{\lambda}$ according to ratio $r = \frac{\hat{I}_1}{\hat{I}_2}$:

$$\hat{\lambda}(N_1, N_2) = -N_1 \ln r + (N_1 + N_2) \ln \left(\frac{N_1 r + N_2}{N_1 + N_2} \right) \quad [8.3]$$

and of studying the distribution of this ratio.

Using the ratio has been proposed mostly for the detection of contours and we will come back to it in the next section. In the case of a multiplying noise, it replaces the difference in average used with an additional noise. The ratio distribution for a single-look image is given, in the most general case, by [OLI 96]:

$$p\left(r \mid c = \frac{R_1}{R_2}, N_1, N_2\right) = \frac{\Gamma(N_1 + N_2)}{\Gamma(N_1)\Gamma(N_2)} \frac{\frac{1}{r} \left(\frac{N_1 r}{N_2 c}\right)^{N_1}}{\left(1 + \frac{N_1 r}{N_2 c}\right)^{N_1 + N_2}} \quad [8.4]$$

For two given thresholds (corresponding to the ratio asymmetry according to whether area 1 or 2 displays the stronger radiometry), it is possible to calculate detection and false alarm probabilities according to the real contrast c between both areas and the number of pixels constituting them, N_1 and N_2 :

$$\mathbf{P}_d(c, N_1, N_2) = \int_0^{t_1} p(r|c, N_1, N_2) dr + \int_{t_2}^{+\infty} p(r|c, N_1, N_2) dr$$

$$\mathbf{P}_{fa}(c, N_1, N_2) = \int_0^{t_1} p(r|c, N_1, N_2) dr + \int_{t_2}^{+\infty} p(r|c, N_1, N_2) dr$$

It is possible to obtain a literal expression, which involves hyper-geometric functions [OLI 96, FJO 99a]:

$$\mathbf{P}_d(c, N_1, N_2) = 1 - \frac{\Gamma(N_1 + N_2)}{N_1 \Gamma(N_1) \Gamma(N_2)} [f(t_2) - f(t_1)]$$

where

$$f(t) = \left(\frac{\beta_2 t}{\beta_1 + \beta_2 t} \right)^{N_1} {}_2F1 \left[N_1, -N_2 + 1; N_1 + 1; \frac{\beta_2 t}{\beta_1 + \beta_2 t} \right]$$

and where $\beta_1 = \frac{R_1}{N_1}$ and $\beta_2 = \frac{R_2}{N_2}$.

As mentioned previously, expressions [8.4] and [8.3] are true only for a speckle totally developed on regions 1 and 2. This hypothesis is no longer verified in the presence of a strong diffuser (then we actually have a Rice law; see Chapter 5). In this case, the only false alarm probability on a homogenous speckle area can be calculated with the help of equation [8.6]. In practice, a very low value of \mathbf{P}_{fa} is set

and threshold t_2 to be applied on r is deduced from ($t_1 = 0$ in the case of shiny targets corresponding to area 1).

It is possible to improve these statistical methods by knowing target distribution, particularly in a military context where target backscattering models are generally available [LON 83, OLI 98].

We have considered in this part the case of single-look intensity image. In the multi-look case, generalized likelihood ratio $\hat{\lambda}$ is obtained by multiplying expression [8.3] by the number of looks L .

8.1.3. Other approaches

The two methods presented below are based on a Bayesian modeling of the problem. Among other techniques that can be used, we mention:

- use of the coefficient of variation, which acquires high values in the case of heterogenous areas and therefore particularly in the presence of a strong backscatterer [CAV 93];
- use of mathematical morphology treatments, able to detect particular shapes, for example, openings or high hats which make it possible to detect the radiometric peaks of the image [CAV 93, SCH 94].

These methods are empirical and they display less reliable results, as they are often marred by artifacts (block phenomena, for example).

8.1.4. Conclusion

The target detection problem remains a complex one, and it can be approached from two different angles. In a military context that uses high-resolution sensors, this problem is often connected with recognizing target models; numerous works have been dedicated to this problem, but not all of them have been published. In a civil context, especially for urban environment detection, the problem remains relatively ill formulated. In fact, backscattering mechanisms in the city are still imperfectly mastered (influence from materials, relative object orientations, resolution, etc.) or are incorrectly integrated in the algorithms (complex laws). The very notion of “target” is not clearly defined: should it be connected only to the high value of backscattered energy, or should it be a characteristic of an object’s presence (knowing that, in relation to its orientation, the emitted signal may be weak)? In order to illustrate our subject, Figure 8.2 shows the backscattered signal of bridges from two different points of view from the same sensor. At times, great differences

are observed between the echoes supplied by the same bridge. Thus, in the two far-right images, a bridge that is clearly visible in the view from above is rather blurred in the low incidence.

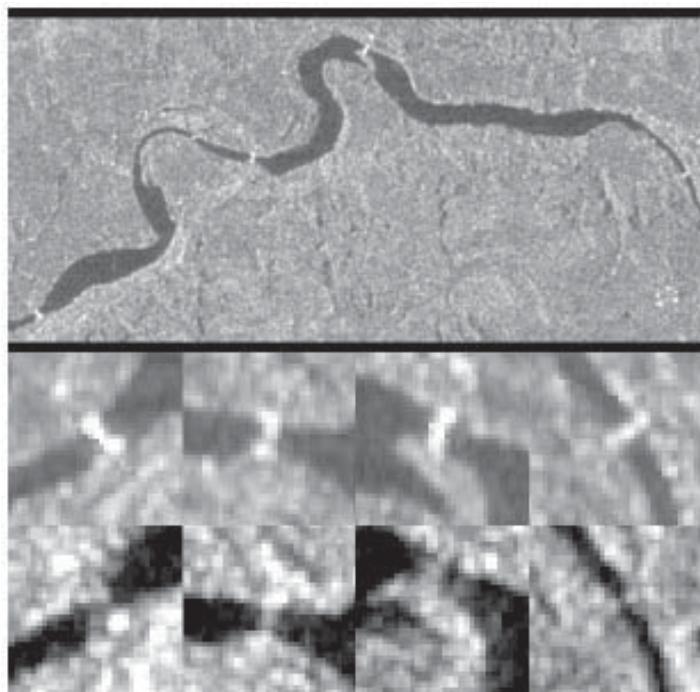


Figure 8.2. Detection of point targets. General view of the area (above) taken by X-SAR©DRL sensor, and enlargement on 4 points situated along the river from two different points of view. It is observed that images of the same bridge appear different in the 2 views (figure supplied by M. Datcu)

8.2. Contour detectors

Although target detection is often used in a particular context, contour detectors are more general tools making it possible to guide ulterior processes (object recognition, surface discrimination, etc.). Their performance conditions the processing chain results in its totality and they are an often crucial low level step.

Most contour detectors traditionally used in image processing are gradient type and therefore based on a local averages difference [LAC 88]. We can show that these detectors are inappropriate for radar imaging because their false alarm

probability P_{fa} increases with radiometry of the areas on which they are applied [TOU 88a]. It would be possible to vary the thresholds used in accordance with average radiometry, but this would greatly complicate their implementation. Therefore, adaptations have been proposed in order to be free of this defect: image logarithmic transformations, average normalization by means of addition [CAV 93], or a function of standard deviations [WHI 91], the use of the coefficient of variation. The difference between image logarithm averages, which is equivalent to the ratio of geometric averages, naturally leads to the use of arithmetic average ratio, which was proposed towards the end of the 1980s by several teams [BOV 88, TOU 88b, MAD 86].

8.2.1. Average ratio

Let us consider first the case of a single vertical contour delimiting two neighborhoods V_1 and V_2 , which are made up of homogenous reflectivity areas R_1 and R_2 , as shown in Figure 8.3 (we will consider the multi-directional case in section 8.2.4). The neighborhood is symmetric when $N_1 = N_2$. The distribution of $r \frac{\hat{I}_1}{\hat{I}_2}$ is always that of equation [8.4], as well as the detection and false alarm probabilities. Contrast c of the contour is defined by $c = \frac{R_1}{R_2}$. Ratio r is the estimator, in the maximum likelihood sense of c .

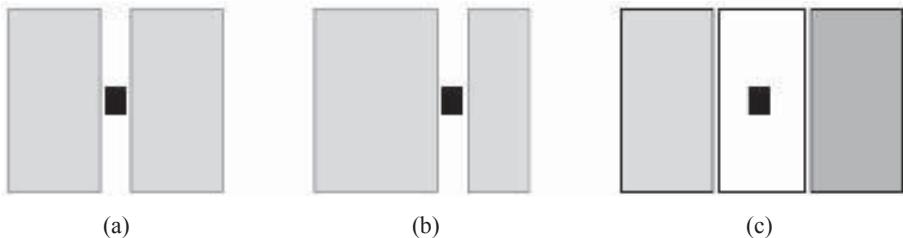


Figure 8.3. Examples of symmetric (a) and asymmetric (b) neighborhoods, which are used for the detection of lines and contours (c)

Since, in general, the relative positions of the shaded and clear areas are difficult to distinguish, the detector can also be defined by [BOV 88]:

$$r_m = \max\left(\frac{\hat{I}_1}{\hat{I}_2}, \frac{\hat{I}_2}{\hat{I}_1}\right) \quad [8.5]$$

In this case, it is also higher than 1 and it increases in accordance with the contour's contrast. In order to have a norm value in $[0, 1]$, we can use [TOU 88a]:

$$r_n = \min\left(\frac{\hat{I}_1}{\hat{I}_2}, \frac{\hat{I}_2}{\hat{I}_1}\right) \quad [8.6]$$

but r_n then decreases in accordance with the contrast; we can also consider $1 - r_n$ and in this case, the distribution is deduced from that of r_n . The approaches by means of equations [8.5] and [8.6] are equivalent and it is enough to change thresholds t_n into $t_m = \frac{1}{t_n}$ in order to pass from one to the other. Distribution r_n is given by:

$$p(r_n) = \frac{\Gamma(N_1 + N_2)}{\Gamma(N_1)\Gamma(N_2)} \left(\frac{N_1}{R_1} \right)^{N_1} \left(\frac{N_2}{R_2} \right)^{N_2} \cdot \left(\frac{r_n^{N_1-1}}{\left[\frac{r_n N_1}{R_1} + \frac{N_2}{R_2} \right]^{N_1+N_2}} + \frac{r_n^{N_2-1}}{\left[\frac{r_n N_2}{R_2} + \frac{N_1}{R_1} \right]^{N_1+N_2}} \right) \quad [8.7]$$

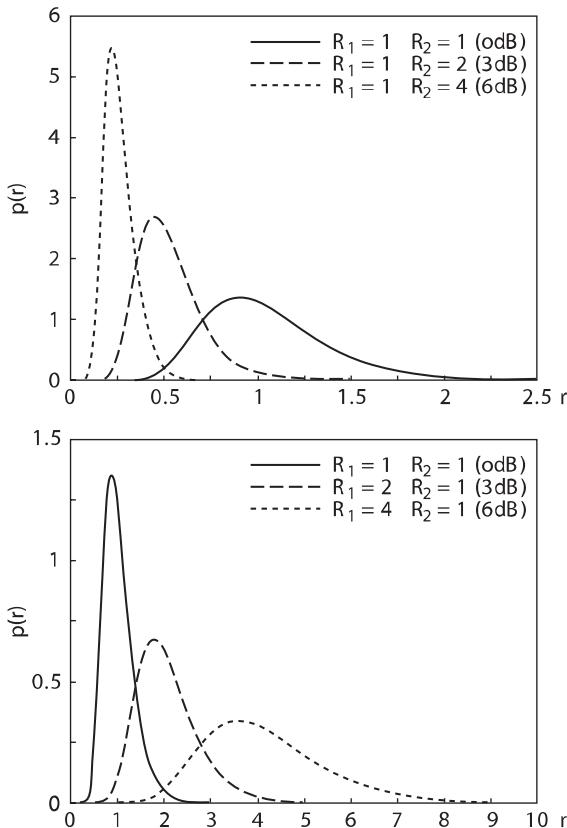


Figure 8.4. Distribution of ratio r across a contour for symmetric windows according to the contour contrast for (a) $R_1 \leq R_2$ and (b) $R_1 \geq R_2$ (according to R. Fjortoft)

Distribution r for various contrasts is represented in Figure 8.4. Distributions r , r_n and r_m are not dependent on $R = R_1 = R_2$ in a homogenous area; this ensures an identical behavior whatever the radiometry of the area (CFAR detector).

In order to calculate probabilities \mathbf{P}_{fa} and \mathbf{P}_{d} corresponding to a given threshold, we can use literal expressions, when they exist, or we may proceed by numerical integration using formulae [8.6] for r , for example. Figure 8.5 shows the detection probability curves according to the false alarm probability (ROC diagram) for various contrasts and window sizes.

Expression [8.7] can also be evaluated in the case of amplitude average ratios. In the multi-look case, the above reasoning and distributions remain valid on condition

that N_i should be replaced by N_iL if the pixels are decorrelated. For correlated pixels, expressions are not valid because the Gamma law needs to be replaced by a Goodman law (see section 5.3.4.2), but if the equivalent number of views L' is used instead of L , very good approximations are obtained [BRU 98]. If complex correlated data exist, it is possible to improve performances, as shown in section 8.2.2.2.

Figure 8.6 shows the application of this contour detector. Because of the size of windows used, multiple responses regarding the exact transition position are obtained and it is therefore necessary to apply post-treatments (see section 8.4).

8.2.2. Likelihood ratio

8.2.2.1. Use of detected data

We address the previous reasoning for target detection by replacing target/background neighborhood with areas located on both sides of the contours. By comparing “contour presence” and “homogenous area” hypotheses, we obtain $\lambda(N_1, N_2, R_1, R_2, R)$, which is given by equation [8.1]. Since area reflectivities are not known, we are limited to using the generalized likelihood ratio, which is written in relation to r , according to equation [8.2].

As mentioned above, it is difficult to establish the distribution of $\hat{\lambda}$. In practice therefore the accepted false alarm probability is fixed and the corresponding threshold is calculated in an iterative manner. Starting from value $t_{\hat{\lambda}}$, which is a threshold on the generalized likelihood ratio, we deduce threshold values t_1 and t_2 on r , which are solutions of equation [8.2]. Therefore, calculation of P_{fa} probability according to equation [8.6] makes it possible to adjust value $t_{\hat{\lambda}}$ up or down and the process is iterated.

The connection between $\hat{\lambda}$ and r shows that $\hat{\lambda}$ is a constant false alarm ratio detector. In the case of a symmetric window for which $N_1 = N_2$, we find t_1 and $t_2 = \frac{1}{t_1}$ as solutions of equation [8.2]; then it is strictly equivalent to use the generalized likelihood ratio or the normalized ratio with corresponding thresholds. In the case of asymmetric windows, it may be slightly more convenient to use $\hat{\lambda}$ rather than r_n , but in practice the results are quite close.

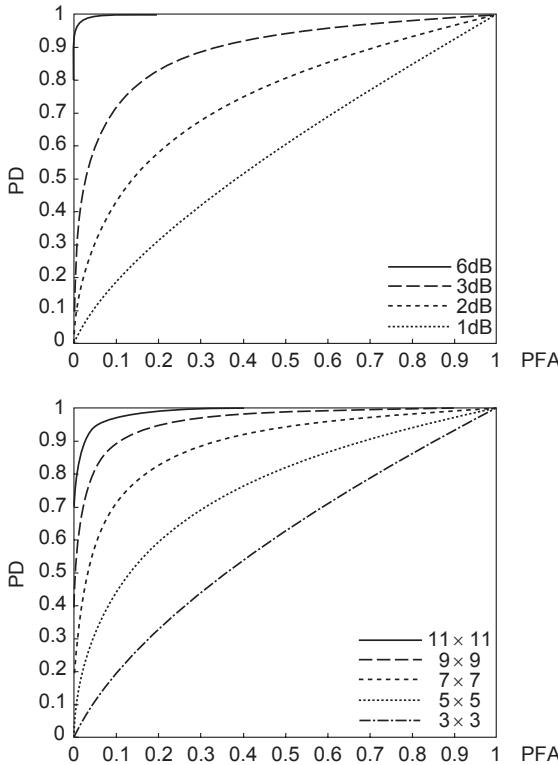


Figure 8.5. Operational curve of the detector ratio receptor. Detection probability of the normalized r_n in relation to false alarm probability for (a) different contrasts and a fixed-size 7×7 centered window and (b) for a fixed 3 dB contrast and various window sizes (according to R. Fjortoft)

8.2.2.2. Complex data use

When complex data exist, it is possible to calculate the likelihood ratio by taking into account pixel correlation. Z_1 and Z_2 are the complex vectors of neighborhoods 1 and 2; we then obtain:

$$f(Z_i | R_i) = \frac{1}{\pi^{N_i} |C_{Z_i}|} \exp(-Z_i^* C_{Z_i}^{-1} Z_i)$$

with C_{Z_i} as the complex covariance matrix of size $N_i \times N_i$, connected to the speckle covariance matrix by $C_{Z_i} = R_i C_{S_i}$. The correlation matrix C_{S_i} is deduced from

speckle covariance matrix C_S by rearrangement of its elements according to the neighborhoods under consideration. By using for R_i the ML estimator:

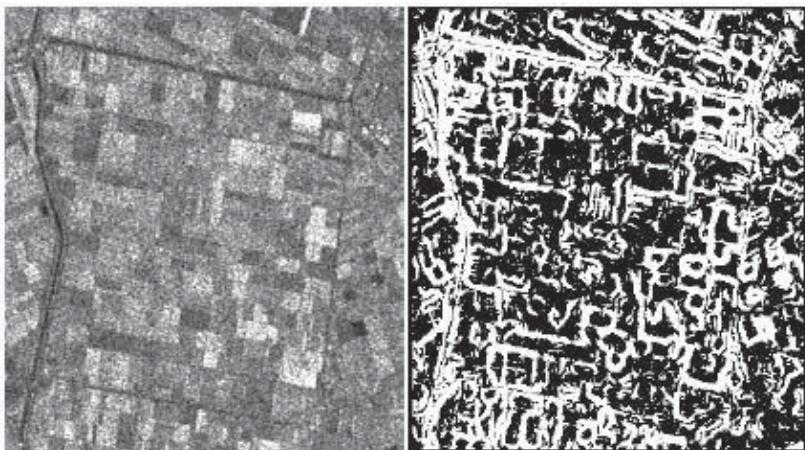
$$\hat{R}_i = \frac{1}{N_1} Z_i^{t^*} C_{S_i}^{-1} Z_i \quad [8.8]$$

we obtain the generalized likelihood ratio:

$$\hat{\lambda} = -N_1 \ln(r) + (N_1 + N_2) \ln\left(\frac{N_1 r + N_2}{N_1 + N_2}\right)$$

with $r = \frac{\hat{R}_1}{\hat{R}_2}$. We find exactly the same expression as above, but with reflectivity ML estimators from equation [8.8], corresponding to complex data. Therefore, this amounts to applying a whitening filter within each analysis window area.

It is also possible to model the speckle correlation with the contour. It is possible to show that the correlation between R_1 and R_2 slightly improves detection performances [FJO 99a].



a. ERS-1 original image from ©ESA

b. Threshold responses

Figure 8.6. Example of contour detection by average ratio on an ERS-1 PRI image

8.2.3. Exponentially-weighted average ratio

The arithmetic averages ratio requires large calculation windows in order to guarantee a reliable detection; this raises a problem when multiple contours are present in the image. Therefore, an improved detector has been proposed in [FJO 98], which uses a non-uniform weighting for the calculation of average intensities. The weighting factors are a distance dependent on the central pixel and they decrease exponentially. This detector is particularly interesting if numerous close contours exist in the image.

The factors used are calculated from the following model, which presupposes the scene to be constituted of homogenous areas. The localization of reflectivity leaps along a certain direction (horizontal or vertical, the model being supposed as separable) follows a Poisson distribution of parameter l , which corresponds to the discontinuity frequency. The probability of k leaps in an interval x , according to a given direction, is written as:

$$p(x) = \frac{lx}{k!} \exp(-lx)$$

From this we can show that the auto-covariance function of reflectivity C_{RR} and its spectral power density S_{RR} , which is its Fourier transform, can be expressed [STE 67] by:

$$C_{RR}(x) = \sigma_R^2 \exp(-lx) \quad [8.9]$$

$$S_{RR}(\omega) = \frac{2l\sigma_R^2}{l^2 + \omega^2} \quad [8.10]$$

Then, if we look for the separable linear filter minimizing the average quadratic error for estimating local reflectivity \hat{R} , in a multi-contour context and with multiplying-noise model, we show that its impulse response has the form:

$$f(x) = C_\alpha \exp(-\alpha x)$$

in the one-dimensional case, with:

$$\alpha^2 = \frac{2Ll}{1 + \left[\frac{E(R)}{\sigma_R} \right]^2} + l^2$$

and C_α as a normalization constant. Parameters are estimated as described in Chapter 6, except for l , which depends on the average length of areas and can be calculated from relation [8.10]. By choosing $C_\alpha = \frac{\alpha}{2}$, we obtain an unbiased estimator and \hat{R} is obtained by the convolution of f and intensity image.

This filter is similar to the Shen and Castan filter [SHE 92], which defines an optimal contour detector, for images cluttered by an additive white noise, by considering the difference of exponentially weighted averages, on both sides, for horizontal and vertical directions. Such a contour detector would be inappropriate in the case of radar images, as already been mentioned at the beginning of this chapter. The idea is therefore to use the smoothing filter, which has been theoretically deduced above, but to replace the difference with a ratio.

The two-dimensional filter is given by $f(x, y) = f(x)f(y)$. This filter is not isotropic. Its behavior responds well to *a priori* elements introduced in the multi-contour model and its response decreases exponentially based on the distance to the pixel processed.

In some cases, f can be quite efficiently implemented with the help of two recursive filters f_1 and f_2 :

$$f_1(n) = a b^n u(n) \text{ and } f_2(n) = a b^{-n} u(-n)$$

with $b = \exp(-a)$, $a = 1 - b$ and $u(n)$ is a discrete Heaviside function. Then smoothing function f can be rewritten:

$$f(n) = \frac{1}{a+b} f_1(n) + \frac{b}{1+b} f_2(n-1)$$

The convolution with f_1 and f_2 is expressed, by noting by S_j the entrance and by S_j' the filter exit, by:

$$\begin{aligned}s_1(n) &= a e_1(n) + b s_1(n-1) \\s_2(n) &= a e_2(n) + b s_2(n+1)\end{aligned}$$

The average estimation on both half-windows on one side and another of the contour is then made, for a vertical contour, by:

$$\begin{aligned}\hat{R}_{x1}(x, y) &= f_1(x) * (f(y) \star I(x, y)) \\ \hat{R}_{x2}(x, y) &= f_2(x) * (f(y) \star I(x, y))\end{aligned}$$

where $*$ symbolizes convolution in the horizontal direction and \star in the vertical direction.

If, as proposed above, we replace the difference by a ratio in order to obtain the contour detector, we also obtain a horizontal gradient estimation:

$$r_x^{\text{exp}}(x, y) = \max \left(\frac{\hat{R}_{x1}(x-1, y)}{\hat{R}_{x2}(x+1, y)}, \frac{\hat{R}_{x2}(x+1, y)}{\hat{R}_{x1}(x-1, y)} \right)$$

In the same manner, we establish vertical gradient formula r_y^{exp} and the response of the contour detector is given by the Euclidian norm of the two gradients.

Figure 8.7 shows the exponential filters used for contour detection. As predicted, the results obtained are higher than those given by the ratio detector in the case of multiple contours. Figure 8.8 illustrates this (the method of contour closure is an algorithm of watershed which gives pixel chains passing through the local maxima; see section 8.4).

The only difficulty in the use of the exponentially-weighted average ratio resides in the estimation of α , but it is possible to take a constant value on any image without marring the results too much.

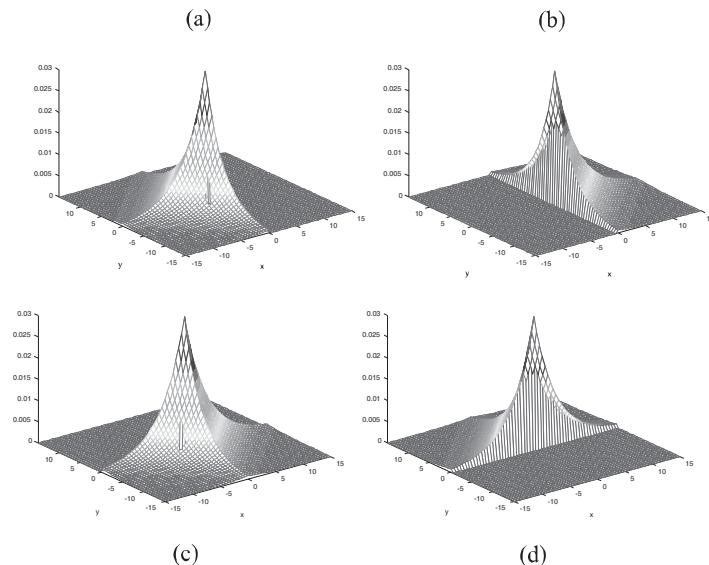


Figure 8.7. Exponential filters used for the detection of vertical and horizontal contours ($\alpha = 0.25$). xy filters calculated respectively: a) \hat{R}_{x2} , b) \hat{R}_{x1} , c) \hat{R}_{y2} , d) \hat{R}_{y1}

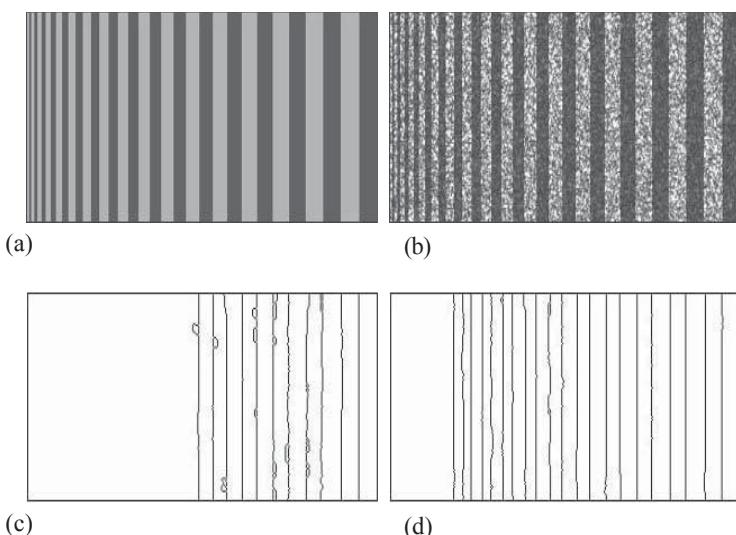


Figure 8.8. a) Original image formed of 2 to 18 pixel vertical lines; b) simulated single-look speckle image; c) segmentation obtained with contour detector ratio; d) segmentation obtained with exponentially-weighted average ratio

8.2.4. Multi-directional case

Until now we have considered the hypothesis of a single vertical direction in order to introduce the detectors presented above. Naturally, in practice it is necessary to take into account several orientations. In this case we use masks in several directions before combining their responses (we will note by \mathcal{N}_d the number of directions).

Two approaches are possible:

- we can use several directions (4 for example: horizontal, vertical and one along each of the two diagonals) and then preserve only the maximal response. This operation increases the false alarm ratio and it is therefore important to take that into account in the choice of thresholds. In this case, the theoretical calculations of distributions are complex because ratios in various directions are not independent, but the following approximation gives good results [TOU 88b]:

$$\mathbf{P}_{\text{fa}}(\mathcal{N}_d) = 1 - (1 - \mathbf{P}_{\text{fa}})^{\alpha}$$

with $\alpha = 3$ when \mathcal{N}_d is 4 (and $\alpha = 5$ for $\mathcal{N}_d = 8$);

- we can also use only two orthogonal directions and combine them in such a manner as to obtain the “gradient norm” according to the following formula:

$$\|\overrightarrow{\text{grad}} g(x, y)\| = \sqrt{\left(\frac{\delta g}{\delta x}\right)^2 + \left(\frac{\delta g}{\delta y}\right)^2}.$$

8.3. Line detectors

In this part we address the detection of linear structures on radar images. In the same manner as target detection, this problem is important because numerous objects, road and rail networks and hydrographic networks appear in this form. We consider here linear structure of width less than 5 pixels. Beyond this value, it is possible to couple detections of contours on both sides of the structure in order to achieve the detection, or to go back to a thinner network using a multi-scale process.

As above, traditional line detectors correspond to radiometric differences [VAN 77, FIS 81], which are inappropriate to radar image treatments and which lead to variable behaviors according to area radiometry. As in the case of contour detectors,

the solution consisted of replacing the difference by a ratio and the accurate statistical modeling of speckle has made it possible to achieve finite studies of probabilities \mathbf{P}_{fa} and \mathbf{P}_{d} .

8.3.1. Average ratio

Let us consider a vertical line and a neighborhood defined as in Figure 8.3c with indices 1 for the central area forming the line, and 2 and 3 for the adjacent areas. Let us note by r_n^{12} and r_n^{13} the normalized detector ratio between areas 1 and 2, and 1 and 3 respectively. If we want to obtain a sufficient contrast on one side or the other of the detected lines, the linear structure detector can be defined by:

$$\eta_l = \min(r_n^{12}, r_n^{13}) \quad [8.11]$$

For a given direction, this comes down to making a contour detection on both sides of the line and preserves only the less satisfactory response.

By using r_n distribution it is possible to deduce r_l distribution. In fact, by noting by Φ the distribution function of the normalized ratio, we obtain the following relation:

$$p(\eta_l) = p(r_n^{12})\Phi(r_n^{13}) + p(r_n^{13})\Phi(r_n^{12}) \quad [8.12]$$

By using the distribution of r_n given by equation [8.7] it is possible to estimate that of r_l in relation to contrasts $c_{12} = \frac{R_1}{R_2}$ and $c_{13} = \frac{R_1}{R_3}$, and in relation to the sizes of windows used, i.e. N_1, N_2 and N_3 , which we will note by $f(r_l | c_{12}, c_{13}, N_1, N_2, N_3)$. Then, for a given threshold s , we have the following detection and false alarm probability:

$$\begin{aligned} \mathbf{P}_{\text{d}}(s, c_{12}, c_{13}, N_1, N_2, N_3) &= \int_s^1 p(\eta_l | c_{12}, c_{13}, N_1, N_2, N_3) d\eta_l \\ \mathbf{P}_{\text{fa}}(s, c, N_1, N_2, N_3) &= \int_s^1 p(\eta_l | c_{12} = 1, c_{13} = c, N_1, N_2, N_3) d\eta_l \\ &= \int_s^1 p(\eta_l | c_{12} = c, c_{13} = 1, N_1, N_2, N_3) d\eta_l \end{aligned}$$

The false alarm probability now corresponds to the probability of falsely detecting a line in a totally homogenous area for $c = 1$, or detecting a line instead of a contour for $c \neq 1$.

Figures 8.9a and 8.9b show examples of P_{fa} and P_d .

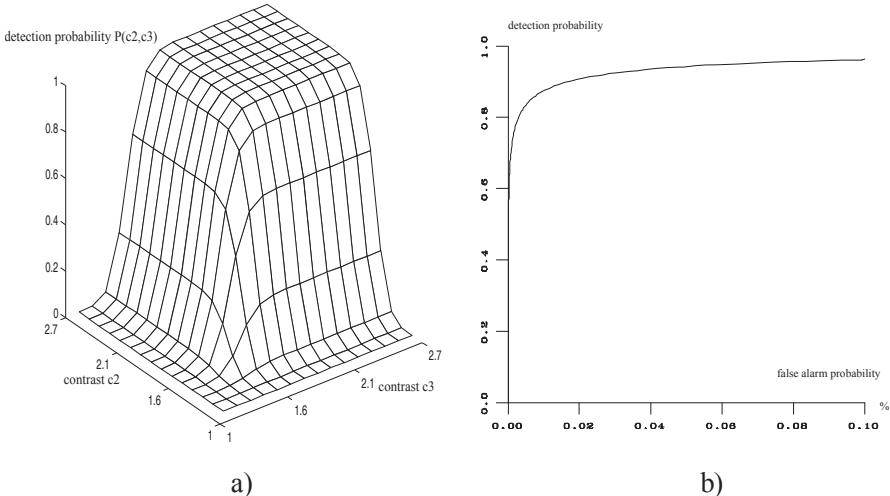


Figure 8.9. Detection probabilities of line detector r : (a) detection probabilities in relation to $\sqrt{c_{12}} = c_2$ and $\sqrt{c_{13}} = c_3$ for $s = 0.3$ and $N_1 = 33$, $N_2 = N_3 = 22$. (b) P_d for $c_{12} = 4$ and $c_{13} = 2.25$ according to P_{fa} for a contour with $c = 4$

8.3.2. Correlation with an ideal structure

Another line detector uses the normalized correlation factor centered between an ideal contour model and the neighborhood, on both sides of the line [TUP 98]. Starting from a given configuration of various neighborhood radiometry data, and supposing the existence of a contour in a fixed direction, the ideal contour model is, in fact, formed of two areas where radiometry is the average of the pixels constituting them. Then we can establish for the correlation detector, noted by ρ , the following expression:

$$\rho^2 = \frac{1}{1 + (N_1 + N_2) \frac{N_1 \gamma_{R_1}^2 r_n^2 + N_2 \gamma_{R_2}^2}{N_1 N_2 (r_n - 1)^2}}$$

The derived line detector is then $\rho_l = \min(\rho^{12}, \rho^{13})$ by taking the same notations as above. The expression of ρ not only involves the contrast, but also considers the homogeneity of both zones. Therefore, there will be a different r_n behavior in non-homogenous areas and there will be a weaker response, particularly when an isolated target comes to disturb the calculation of the averages (see Figure 8.10).

Since it is not possible to establish distribution of ρ , the statistical study of this detector can be achieved only by simulation.

Rather than choosing one of the two detectors r_l or ρ_l , it can prove more interesting to merge their responses. Several fusion operators of line detectors in a multi-temporal framework are studied [CHA 99].

8.4. Line and contour connection

As with all detection methods of contours and lines, low level detectors are followed by post-treatment operations making it possible to refine and complete contours. This is particularly true in the case of radar images for which large-size masks are used. It is then necessary to refine the contours either by searching for the maximal response position in the direction perpendicular to the contour, or by mathematical morphology operations on the threshold result [TOU 88b].

Concerning contour closure, segmentation logic is often used. For example, it is possible to achieve an increase of areas conditional to contours [WHI 91]. Let us also cite the use of the watershed algorithm, making it possible to obtain an over-segmented image containing most of the true contours. Then the number of false contours is reduced by an area fusion stage [FJO 98] (a segmentation result is presented in Chapter 7). A repositioning of borders by likelihood and regularization maximization can also be achieved with Markov fields or active contours on the image of vector contours [FJO 99b, GER 98].

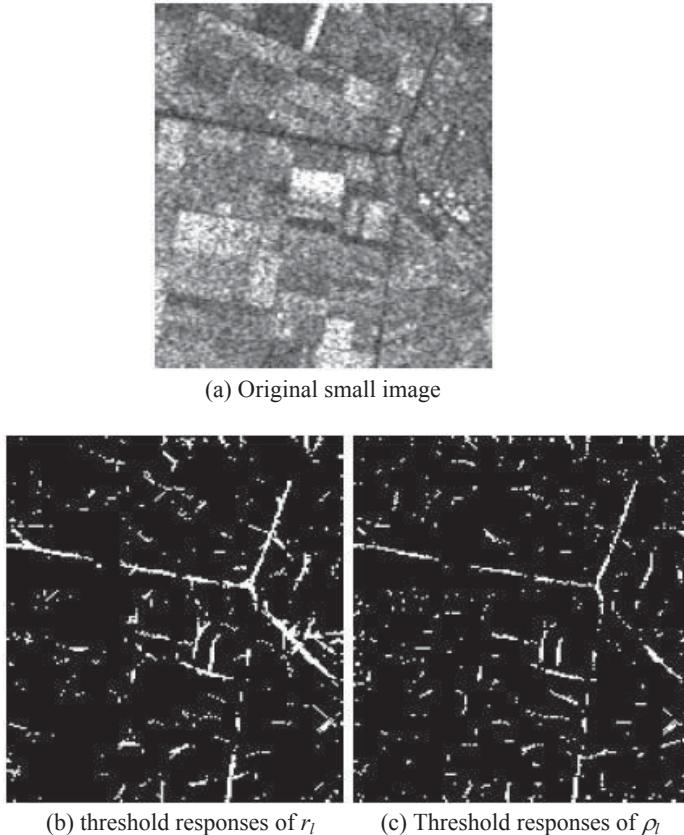


Figure 8.10. Illustration of line detectors r_l and ρ_l on a small image extracted from an agricultural scene from the north of the Netherlands ERS-1 ©ESA

Concerning line detection methods, and in order to improve on the relatively mediocre performances of low-level detectors on radar images, more elaborate techniques have emerged. *A priori* information can be thus introduced inside the line detection process [HEL 96b, TUP 98]. Interactions between detected contours are modeled, especially by imposing a certain contour regularity. Thus, the contour pixels whose neighbors are also contours with similar directions will be favored. Since such types of constraints are essentially of contextual origin, Markov fields are a particularly adapted framework and they are often used for this purpose. Figure 8.11 illustrates this type of method in the case of road network detection (global area cartography is obtained by the method described in [TUP 99]; see also the classification in Chapter 7).

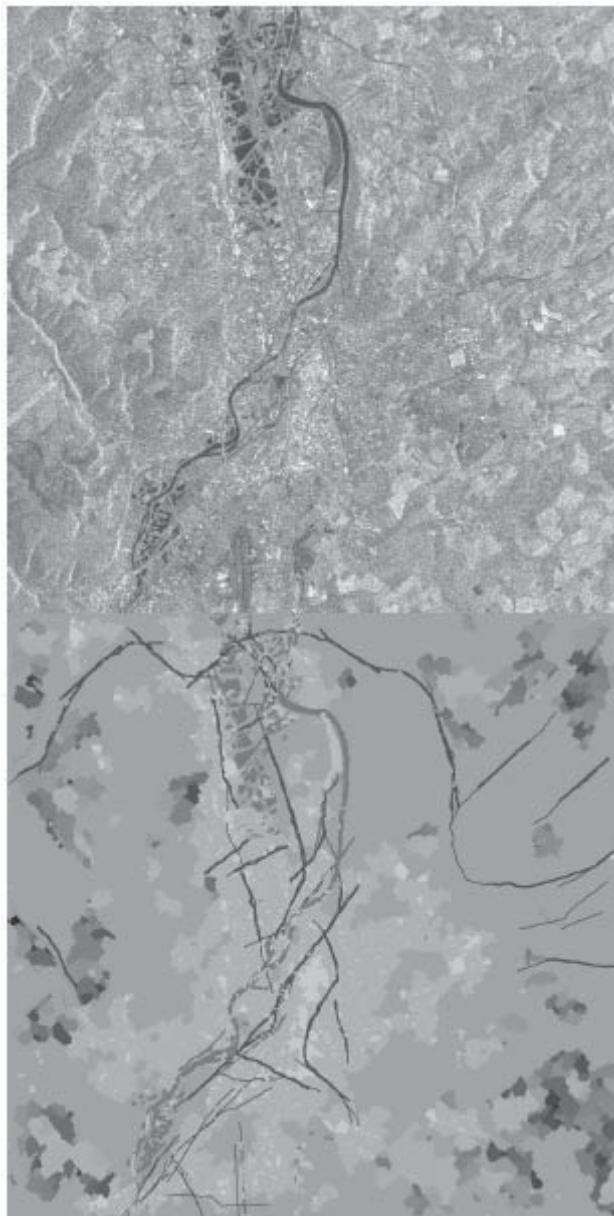


Figure 8.11. Illustration of network detection on an ERS-1 ©ESA image focused on the Metz area; the cartography set makes it possible to classify the following objects: in black, rivers and forests; in gray, the road network; in white, the decreasingly dense urban environment) (see color plate section)

8.5. Conclusion

As with filtering, most structure detection methods, whether they are targets, contours, or lines, are based on statistical analysis leading to adaptable contextual operators optimized for the specific speckle properties. Generally, these detectors are based on generalized average or likelihood ratios; the operators are based on ML estimations of unknown parameters.

We have presented here the most commonly used techniques. Certainly, other approaches have been proposed; however, as in the case of filtering, it is difficult to make rigorous evaluations of various detectors. In fact, performances are often conditioned by the type of data to which they are applied (high-resolution or satellite) and by the final application considered. In practice, following the post-treatments or subsequent chain steps, it will be advisable to favor a strong detection rate or a weak false alarm rate.

However, raw performances are well below those obtained with optical images because of the limits of local detectors and data difficulty, which is particularly strong in the case of single-look speckle and a significant pixel correlation. The use of SLC data makes it possible to reduce, or even to eliminate, the negative effect of strong correlation. Since multi-scale approaches make it possible to make detection more robust [TOU 88b] and post-treatments more elaborate, they are often applied to improve results. These post-treatments should complete detection in a coherent manner and therefore they should be adapted to radar imaging.

This page intentionally left blank

Chapter 9

Geometry and Relief

We have seen that SAR images could be produced in radar or ground geometry: the approach that has made it possible for us to make this distinction was based on hypotheses of flat land as well as on perfect sensor knowledge. In this chapter, we will approach issues related to sensor localization and image registration.

9.1. Radar image localization

The main geometric difference existing between radar imaging and traditional (optical) imaging is due to the fact that it only involves distances and not angles. Due to this fact, geometric modeling of a radar image is clearly simplified as compared to a conventional image as produced by a lens. In the same manner, localization precision will be linked only to satellite position and not to its attitude¹.

Chapter written by Frédéric ADRAGNA.

¹ The term “attitude” refers to the satellite’s orientation on its orbit. It is located by means of 3 angles called yaw, roll and pitch.

9.1.1. Precision of orbital data

It is necessary to differentiate between three precision types, which in fact, correspond to the three photographic stages:

- *a priori* precision: this is the precision with which we know the satellite's position when it is taking the image (programming precision);
- instant precision: at the moment of shooting, the satellite (if it is equipped with navigation instruments) knows its position and can include it in the auxiliary data attached to the image file;
- *a posteriori* precision: a few days after acquisition, this position will be refined by calculations using orbital models (terrestrial potential, friction, etc.) and position measurements previous and subsequent to the acquisition. Then we speak of returned orbits.

It is difficult to offer a rule determining these precisions because the current systems are quite different and techniques evolve according to positioning systems such as GPS or Doris, but we can say that there is an order of magnitude when we pass from one precision to another: the precision is decametric for *a priori* precisions, it is metric for instant precisions and decimetric for *a posteriori* precisions.

It is also necessary to differentiate between radial (vertical), lateral and longitudinal precisions. In this case as well, it is possible to simplify by saying that there is approximately one order of magnitude between these two. For example, for ERS and for restored orbits: the precision is centimetric on altitude, decimetric in lateral and metric in longitudinal.

9.1.2. Distance localization

The useful parameters for radar image distance localization are satellite position (its altitude and its radial and lateral position), proximal distance (distance to the close image boundary) and ground altitude.

A satellite lateral localization error has the effect of shifting the entire image laterally. Since precision of restored orbits is metric (or even better sometimes) in practice, this error does not generate a significant error.

An error regarding effective satellite altitude (or, similarly, regarding ground elevation) generates a distance localization error and also a distortion (the shift depends on the incidence). If the satellite is lower (or if the ground is higher,

respectively), the image is situated farther from the trace, according to the following relation:

$$\Delta d = \frac{\Delta h}{\tan(\theta)}$$

Thus, a 1 m altitude error on an ERS orbit creates a 3 m horizontal shift to the proximal distance and a 2 m shift to the far range (shift and dilation addition).

An error on proximal distance may appear because of an electronic delay variation within the radar (for example, a precision in the 10 m order is generally admitted because of temperature variations) or because of the dilation of the atmosphere being crossed (the total elongation is of several meters, with variations of some tens of centimeters). The elongation caused by crossing the troposphere depends on temperature, pressure and humidity conditions, and it is independent of wavelength. Elongation caused by passing through the ionosphere depends on CET (see section 1.1.3) and wavelength². It generates a horizontal image shift according to the following relation:

$$\Delta d = \frac{\Delta h}{\sin(\theta)}$$

as well as image distortion.

9.1.3. Azimuth localization

Parameters useful for radar image azimuth localization are satellite position (in the longitudinal sense), dating of the given image and the Doppler effect of the photograph.

Certainly, the satellite position in a longitudinal sense (along the orbit) determines the image position in this direction. This position is known (*a posteriori*) by only a few meters and it is given in relation to a clock connected to the platform (station update and satellite orbital monitoring). Raw radar image data (echoes) are also accurately dated, but this is done in accordance with a clock connected to the instrument and therefore within the satellite's payload. Curiously, however, the two clocks are not always synchronized. This is caused by the separation of jobs: for

² These dependencies will be discussed at length in section 12.5.3.

example, in the case of ERS satellites, orbital monitoring is ensured by ESOC (in Germany), while all the products associated with image formation are handled by ESRIN (in Italy). These clock shifts (with their variations, corrections, etc.) can generate azimuth localization errors.

It is important to analyze the Doppler effect on azimuth localization. We have seen in Chapter 2 that radar image synthesis is the focusing of a bandwidth created by the relative movement between satellite and points lit by the antenna diagram. In the spatial case, because of the Earth's rotation, the focus perpendicular to the orbit is not at "zero Doppler". Figure 9.1 shows the shift existing in the equator.

In order to compensate for this effect, the ERS platform is "yaw navigated". This rotation, oriented forward in upward direction and backward in downward direction, is maximal at the equator (about 3°) and zero at the poles. Thus, the Doppler photography effects are always close to zero (in practice, they are between -0.3 and $+0.3$ PRF). The RADARSAT platform does not have this advantage. Doppler photographs vary more in accordance with altitude and incidence (± 12 PRF), which consequently complicates processes.

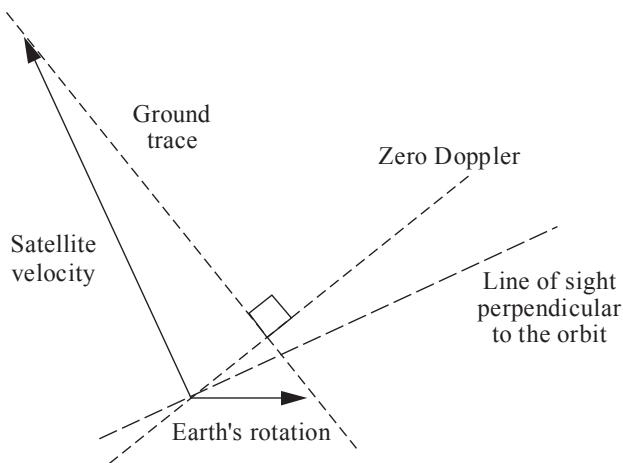


Figure 9.1. The Earth's rotation creates a non-zero Doppler effect, here represented at the equator

From a geometric point of view, this does not change anything, as long as we are able to know where the zero frequency is situated in relation to the sampled frequency band. For ERS, this frequency is always contained in the spectrum and it does not yield any possible error: a centroid Doppler estimation error would entail only a radiometric alteration of the image (bad optimization of useful bandwidth),

and no geometric or localization error. For RADARSAT, it is first necessary to remove the ambiguity connected to the fact that we only have at our disposal the folded spectrum (known by a full *PRF* number). This ambiguity removal is automatic because of the platform's altitude data. Unfortunately, the formula leaves room for a slight uncertainty because there can be estimation errors with platform misalignment.

From the above it is possible to deduce that radar images undergo very few geometric distortions (other than those due to relief) and that they have the advantage of a very good absolute localization. Taking support points, for a finer image registration with a map, will essentially have the purpose of determining the two parameters: a longitudinal (azimuth) shift and a radial (distance) shift.

9.1.4. Three-dimensional point localization from radar imaging

Radar imaging gives the possibility of precisely locating a point on a circle defined by a plane perpendicular to the carrier's trajectory and radial distance to the carrier. In order to situate the point in space (in 3D), we can use an image obtained in orbit of the opposite direction (ascending or descending). Then the point is placed at the intersection of two circles as indicated in Figure 9.2. This property will be developed in Chapter 10, which will deal with radargrammetry.

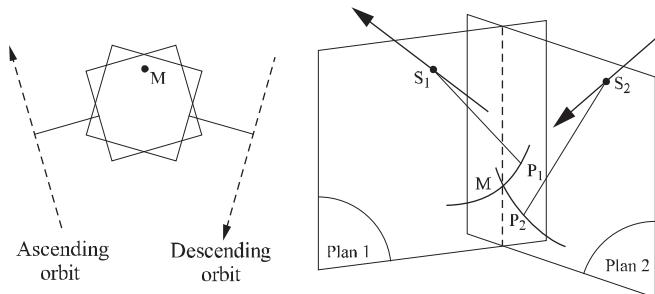


Figure 9.2. Point localization starting from two opposite views.
Left: projection on Earth surface; right: three-dimensional representation

The precision may be quite high (less than 20 m horizontally and 5 m vertically for ERS), but the difficulty lies in locating and recognizing the same point in both images. In particular, it is necessary to be careful of buildings; we can never know exactly which part of these illuminates the radar image (especially with the opposite incidences); we should be careful of crossroads often with trees, bridges, etc. The best way is to choose a characteristic point, which is easy to locate, on flat well-

cleared ground. A study has shown that excellent results can be obtained by using crossroads of airport runways.

9.1.5. Moving targets

It is necessary to emphasize a well-known characteristic of radar imaging, namely shifting targets in movement. We must distinguish between targets with radial movement and those with a movement parallel to the radar. In this latter case, the target will simply be badly focused as this comes down to incorrectly estimating the (relative) velocity of the sensor compared to this target during synthesis. In the case of radial movement, Doppler frequencies are shifted generating an azimuth movement. The shift is given by formula:

$$\delta = \frac{V_r}{V_s} D$$

where V_r is the target's radial velocity towards the radar, V_s is the satellite velocity and D the radar-target distance. Figure 9.3 shows that this shift can be justified geometrically without using frequency formalism. Ship *A* advancing towards the radar while photography is taken is always the same distance from the radar as compared with ship *B*, which is stationary. Therefore, the two echoes will be focused in the same position, namely that of *B*.

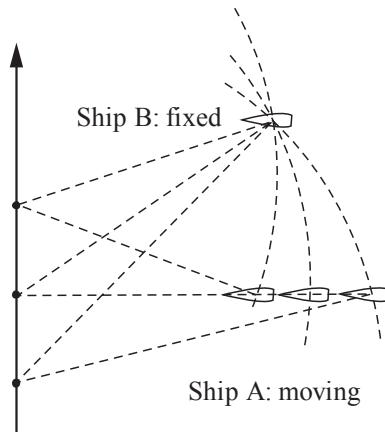


Figure 9.3. A moving object *A* in the scene will move during image formation and will be interpreted as a stationary target placed at point *B*

9.2. Geometric corrections

9.2.1. Geometric corrections with DTM

Before correcting the image and “flattening” it on a DTM, it is necessary to refine its absolute location. We have seen in section 9.1 that each pixel can be perfectly localized with only the following knowledge: orbit, time, initiation of image acquisition, proximal distance used and a digital ground model. Any other parameters, such as satellite altitude, focal plane deformation, optical (instrument) defects or possible vibrations should not be taken into account. It is here that one of the advantages of radar imaging lies compared with other observation systems such as optical sensors.

In practice, however, each of these parameters is known only with approximate certainty (orbital deviation, dating imprecision, dilation of lengths when crossing the atmosphere, electronic delay, etc.). Before all geometric treatment, it is necessary to remove these uncertainties. It is demonstrated that if the DTM is taken as cartographic reference, it is then possible to suppose the exact orbit and to bring back the totality of the problem to the determination of two parameters: acquisition initiation time and proximal distance [SYL 97].

Any lateral or vertical shifting of the orbit will be included in the proximal distance uncertainty. Any longitudinal bias of the orbit will be included in the uncertainty related to the acquisition initiation time.

In order to determine these two parameters (simple translation in radar geometry), it is possible to take support points, namely points identified in the image and also observed in a detailed map in the same cartographic system as the DTM. In this case, a few points are enough to deduce a translation from setting in correspondence the 2 pairs of points. However, in general, more points are considered ($n = 5$ to 20) and a system of n equations with two unknowns is solved by least squares estimates. A solution that is sometimes preferred consists of simulating a radar image starting from the knowledge of the orbit and DTM, then correlating this image with the measured radar image. The similarity of simulated images is generally mediocre (except for the case when there is much relief), but then a fairly over-dimensioned system is solved, which generally guarantees a quality solution. Correlation then reaches excellent precision rates (typically lower than 10 m).

Once image registration between the image and DTM is achieved, we have all the parameters for re-sampling the image and transferring it on the DTM. The method consists of scanning the DTM (re-sampled at the desired resolution) and, for each point, finding its counterpart in the radar image (*reverse projection principle*).

The approach that would search the image in the DTM from each radar image point (called *direct projection*) is not possible because a pixel of the image may correspond to several points of the land in the case of folding. If image rendition by means of SAR synthesis has been carried out at zero Doppler (namely on a plane that is perpendicular to the orbit), this considerably simplifies the process. For each DTM node, this process consists of:

- determining its projection perpendicular to the orbit (in an analytical manner or, more simply, by iteration), which gives the time for the acquisition of this point and consequently the corresponding radar image line number;
- determining the distance between this point and the orbit. After subtracting the proximal distance, we deduce from it the column number in the radar image.

After that, the choice of the best attribution strategy for a radiometric value (closest neighbor, local re-sampling, filtering, smoothing, compensation of slope effects, etc.) is widely open. This choice depends on the user and need.

9.2.2. *Image registration*

One of the advantages of radar is being able to use significant temporal sequences since meteorological conditions do not affect visibility. This compensates in part the image readability defect due to speckle presence. This also facilitates interferometry and radargrammetry techniques.

Most applications require image registration. In particular, those using segmentation (such as, for example, agricultural classification which uses multi-date series for the survey and classification of cultures [BRU 94, LOP 97b]), pattern recognition (for cartographic, national planning [RUD 96]), and certainly interferometry (see Chapter 12) and radargrammetry (see Chapter 10).

However, it is necessary to distinguish between various cases, according to whether orbits of the images under consideration are identical (the same incidence) or not, or even to consider the combination of ascending or descending orbits.

9.2.2.1. *Case of images obtained from identical orbits*

This is the simplest case because alterations from one image to another are quite insignificant and almost never depend on relief. However, it is necessary to distinguish two cases:

- if orbital deviations are low and we are under interferometric conditions. This is the most favorable case because speckle noise is the same in both images and it is going to take part in the correlation, even for uniform areas (absence of contrast).

Correlation between images makes it possible for us to generate a deformation grid of one of the two images, chosen as the slave, and in relation to another one, chosen as the master or reference. Since correlation does not necessarily make it possible to obtain correspondences on any image (presence of low-contrast areas or weak coherence areas), it will be necessary to smooth this grid or to interpolate a simple polynomial model. Better still, if we have a digital model of the land at our disposal (even rough), it is possible to predict a theoretical alteration grid and register it using the measured grid. This technique (which will be also used in radargrammetry (see Chapter 10) and in interferometry (see section 9.2.1)) produces excellent results;

- if the images are not in interferometric conditions (more significant orbital or temporal deviations). The problem is the same, but operations run the risk of being less accurate because of a more unsatisfactory correlation quality. Fortunately, distortions between the two images are weak and the deformation model is generally simple (it requires a smaller number of areas that “connect”). This is the case of multi-temporal fusion for radiometric improvement.

9.2.2.2. Images acquired from different incidences

If we do not have a digital ground model and if we wish to build it, we can use radargrammetry techniques [SYL 98] discussed in Chapter 10. If the objective is to register several images with the help of a DTM, the problem comes down to the case in section 9.2.1, in which each image is handled separately and projected on the same DTM.

At this stage, it is important to note that, since radar imaging radiometry is quite sensitive to incidence, it is quite difficult to compare or to merge images acquired on various incidences. The majority of applications using several radar images (temporal series, comparison before/after an event) require working on a single incidence. Certainly, this limits the advantage of multi-incidence systems. Using these multi-incidence data clearly makes it possible to gain access to information in time, but this is compensated by operation problems. However, multi-incidence imaging appears to be of great interest for problems in which geometric information is paramount. This is particularly the case of urban environments and especially of very high resolution imaging, which makes it possible to provide details on city structures. Multiplying photographic orientations is therefore fundamental to compensate for the anisotropy resulting from façade or road alignments, for example [TUP 00, BIA 99].

9.2.2.3. Images obtained in opposite direction orbits (ascending/descending)

In most cases, it is not possible to correlate the images obtained in the opposite view (ascending orbits/descending orbits or left view/right view). On the one hand, image radiometry is sensitive to incidence, and here we have an extreme case. On

the other hand, the objects (buildings, trees, mountains, etc.) are seen from their opposite face and they appear quite differently.

Essentially, the applications of ascending and descending orbit crossing are geometric image registration and support point of reference by spatial triangulation, which is detailed in section 9.1.4.

Chapter 10

Radargrammetry

Stereoscopy is the most natural way for human beings to perceive reliefs; stereoscopy is the use of distance transformation properties by projections on two sensors placed at different positions, whether these are the eyes of an observer or photographic receptors. On that account, radar imaging will enable stereoscopic reconstruction. However, we will see that stereoscopy, sometimes called *radargrammetry*, just as stereovision applied to landscape readings is called photogrammetry, is different from the latter in various ways.

Radargrammetry emerges as one technique at our disposal for reconstructing reliefs from radar images [POL 97, TOU 00]. This is one of the simplest methods. However, it requires significant adjustments of the diagram used in optical imaging (see section 10.1.2). These adjustments will be examined in section 10.2.

10.1. Stereovision principles: photogrammetry

10.1.1. *Principles*

In the domain of optical wavelengths, remote sensing applications have often used stereovision techniques in order to reconstruct three-dimensional information, either landscapes or human constructions [AYA 89, HOR 93].

Reconstruction systems (called stereoscopic plotters), analog at first, then hybrid and finally entirely digital, are commonly used by all cartographic agents and make it possible to obtain remarkably accurate and reliable altitude charts.

The principle of these reconstructions always follows the same steps:

- first, a viewing system is placed in two different positions and registers two images of the scene to be reconstructed; this viewing system may be a satellite in the course of two passages above an area or an airborne platform during a systematic mission. During each viewing, the position of the sensor and its orientation are stored with great precision. These images are traditionally called “right image” and “left image”, with reference to the human visual system;

- then we determine, for each image, the geometric transformations which associate land coordinates to image coordinates. These transformations are generally functions because numerous points in the scene can potentially correspond to a point in the image (an entire straight line in most photographic optical systems);

- we then search for point pairs, i.e. scene images of the same point, on the right and left images. This type of research, called matching or corresponding, is either carried out systematically on all the scene's visible points, or on a certain selected number of points;

- using the geometric transformations determined previously, we reconstruct the three-dimensional scene; this means that we calculate, for each image point pair, the single 3D point satisfying both projection equations, or, if no exact solution exists because of noise, the single 3D point minimizing the error. This is the 3D reconstruction step;

- this step is accompanied, or followed, by a “regularization” stage, which ensures that the reconstructed 3D points are a globally acceptable solution; this means that the reconstructed scene should possess the expected geometric properties (continuity and regularity of surfaces, including hidden parts depending on depth of field, respect for known dips, etc.)

10.1.2. Stereovision methods

In optical imaging and particularly for remote sensing applications, all the previous steps have received great attention and are now well mastered:

- photographs are often taken in vertical or quasi-vertical views (this reduces hidden parts), with overlapping from 25% to 80% between successive images. The ratio between stereoscopic base (distance between optical centers during two viewings) and land distance varies from about 0.25 to 2;

– the geometric transformation associated with a no-distortion standard camera is the homographic projective transformation. It is determined from the camera's 3 optical center position parameters, from the 3 optical axis orientation angles and from the focal distance (external orientation of the camera), as well as from the two coordinates of the optical center projection onto the image plane and from the two basic sampling vectors (internal orientation). Homography associates a straight line from real space (viewing line) to an image point. By using homography properties, it is possible to simplify the search for a corresponding point because it can be shown that the corresponding image point is necessarily found on a straight line (epipolar point line) of the other image. In order to take advantage of this property, we often represent images in epipolar geometry, in which the corresponding points appear on lines of the same level in each image. However, this is accompanied by a slight image degradation because it is necessary to resample it. Other satellite sensors (such as *push-brooms*), have different transformation equations. They do not always offer the advantage of accurate epipolar geometry. In this case we apply approximations of epipolar geometry in the areas where altitude variation is low;

– matching is most often achieved by correlation techniques. A small window is defined around the point where a correspondent is sought. An identical window is transferred to the other image, along the epipolar line associated with the point in each position. The correlation between both windows is measured; this correlation is generally averaged and centered. The best correlation indicates the best candidate to matching. Dynamic programming techniques or hypotheses tests are also used sometimes;

– from the knowledge of two homologous points, we find the 3D space point position as an intersection of the two viewing lines;

– the constraints most often associated with reconstruction are surface “almost everywhere” continuity constraints and depth continuity constraints along surface discontinuity lines. Sometimes we add sequence constraints between various objects from both looks, which translate the opacity of objects from the observed world and surface uniqueness. It is often at this stage that adequate solutions are found for the difficult problem of parts seen in one scene but hidden in another.

10.2. Principles of radargrammetry

The efficiency of photogrammetry techniques has suggested quite early the development of similar methods in radar imaging. We already find the formalism of radargrammetry in [LAP 63]. Later, research on this theme was greatly developed by Leberl [LEB 86, LEB 90].

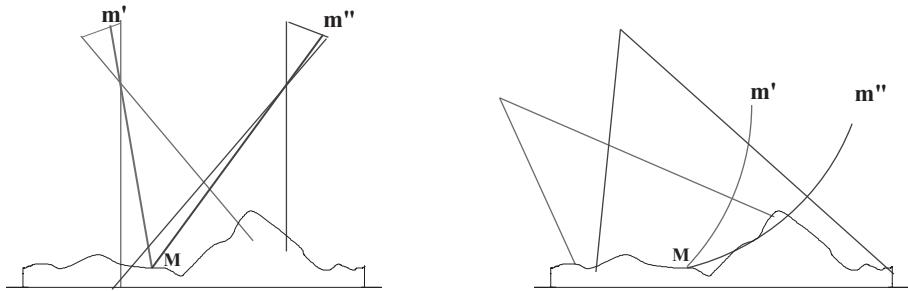


Figure 10.1. Stereovision principle in optics by ray-tracing (left) and in imaging by time-of-flight (right). In the first case, a point M is reconstructed at the intersection of both views passing by the two image points m' and m'' . In the second case, point M is at the intersection of two circles the radii of which correspond to each signal's time-of-flight

However, stereovision techniques cannot be so easily transposed in radar imaging for various reasons:

- first, viewing by “lateral vision” induces certain constraints on image acquisition, which optical imaging does not generally have;
- then, the “time-of-flight” at the basis of geometric coding in radar imaging imposes certain geometric relations; these relations intervene sometimes in the course of epipolar geometry determination, but especially during reconstruction;
- then, radar image formation, which strongly reflects surface orientation properties in relation to the transmitter, introduces errors and limitations that are unknown in optics;
- finally, speckle strongly disturbs the matching operation.

10.2.1. Lateral vision effects

Due to the viewing conditions by lateral vision on the one hand and by opening synthesis on the other, the RSO radar image is located by two specific orientations: of azimuths and of distances (see section 2.1), which are established by the carrier’s trajectory. Therefore, the carriers’ trajectory defines stereoscopic acquisition properties, as represented in Figure 10.2.

Configurations where trajectories are indifferent (Figure 10.2d) are very seldom used. They lead to rather complex geometric transformations (dealt with in [LEB 90]), where the stereoscopic basis evolves within the same image and which require

good knowledge of the carriers' trajectories. In addition, they raise serious matching problems, which will not be discussed here.

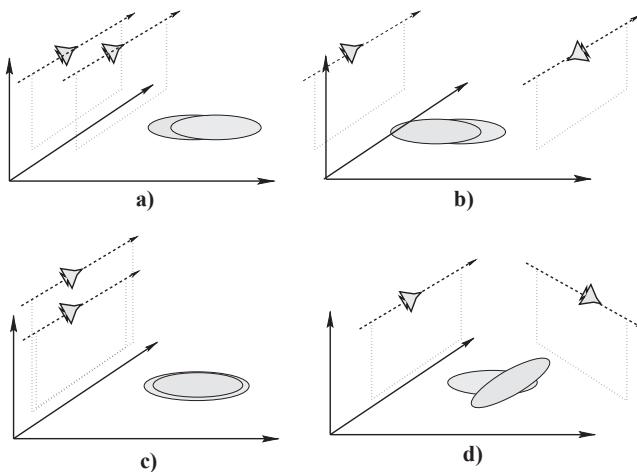


Figure 10.2. Four possible radargrammetry configurations: configurations a) and b) correspond to the most standard ones (particularly in satellite imaging), where trajectories are parallel; in a) the two radars are on the same side, in b) the radars are in opposed positions, in c) the trajectories are parallel but at different heights and superimposed and in d) the trajectories are non-parallel

On the contrary, families of parallel trajectories have generated quite detailed studies because, on the one hand, their implementation is much simpler and, on the other hand they are often the only ones available (for example, in satellite imaging because a single point is generally covered during quasi-parallel trajectories by a passing satellite).

There are two families of parallel trajectories: a family for which the two scene views are on the same side of a vertical plane parallel to the trajectory (Figure 10.2a and c), and a family for which these points are on both sides (Figure 10.2b). When the radar is fixed to the carrier, these configurations correspond to trajectories sometimes in the same direction (a and c) and sometimes in the opposed direction (b). In satellite imaging, as we have seen in Chapter 3, the satellite's direction is described by the terms "ascending orbit" and "descending orbit".

Configuration (c), which requires a different altitude between the two data entries, is generally only accessible with airborne platforms. In addition, it has been shown that it did not present many advantages from the measurement precision standpoint [ROS 68]. Configuration (a), for fixed-angle satellites, only enables low

stereoscopic bases because the maximum accessible angle corresponds to a point that would be at the border of the western field in one view and at the border of the eastern field in the other view. In ERS for example, this leads to a maximum basis/height ratio lower than 0.1, which cannot lead to very high precisions. Variable angle satellites, such as RADARSAT, have a much larger latitude of stereoscopic base choice. This has been discussed in Chapter 3 in section 3.2.4 and is illustrated in Figure 3.4.

Opposite trajectory configurations sometimes lead to strong stereoscopic bases because the angle between the two data entries is twice the angle of the satellite. Thus, in the case of ERS, a radar with low lateral vision (incidence angle of 23°), the basis/height ratio already has a value of 0.85. We will see that this poses difficult matching problems. However, it is these configurations that are the most researched, such as in optics, because they enable a better compromise between “spatial resolution/altimetric resolution”, when the angle between the views is of the order of 90° [ROS 68].

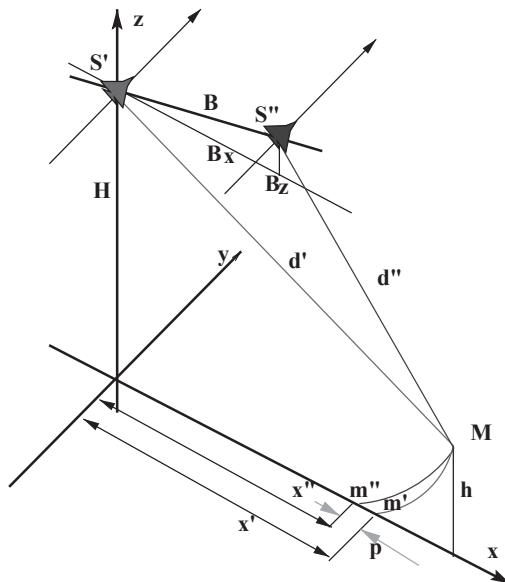


Figure 10.3. The two radar positions are detected by S' and S'' , its height in relation with the reference plane is H , the stereoscopic basis is B , which is decomposed into a horizontal basis B_x and a vertical basis B_z . A point on the ground M , with a height h , has its image in m' and m'' . Disparity is measured by the distance $p = m'm''$

10.2.2. Geometric reconstruction

Below we will only focus on parallel trajectory situations, which are the most common. The general case is presented in [LEB 90]. These trajectories lead to configurations where a form of epipolar geometry can be defined because an observed point to be moved along a vertical line would have, on both images, an image that would migrate along a distance line. Therefore, the radar image is naturally in epipolar geometry if the carriers follow parallel trajectories.

10.2.2.1. Disparity/altitude relation in the general case

Consider Figure 10.3. Point M is seen in m' and m'' in both images. We call *disparity*¹ (and we will note by p) distance $m'm''$. This is zero for a point on the ground ($h = 0$) and it increments for increasing heights h . It is expressed by:

$$p = \sqrt{x^2 + (H-h)^2 - H^2} - \sqrt{(x-B_x)^2 + (H+B_z-h)^2 - (H+B_z)^2} - B_x \quad [10.1]$$

for a given (non-zero) height h , disparity depends on distance x .

More important for the user is the reverse relation which gives height h for a disparity point p . In general, this relation is expressed by an exact but complex form (root of a fourth-degree equation):

$$h = \frac{2H(B_x + p)^2 - (p^4 + 4\Delta)^{1/2}(B_x + p) + 2B_x B_z(x + p) + B_z p^2}{2(B_x^2 - B_z^2 + 2pB_x + p^2)} \quad [10.2]$$

with:

$$\begin{aligned} \Delta = & -x^2 p^2 + p^3 B_x + H^2 (B_x + p)^2 + B_x^2 p^2 + H B_z p^2 \\ & + x^2 B_z^2 + x B_x p (B_x - 2x + p) + 2H B_x B_z (x + p) \end{aligned}$$

¹ There are several definitions of disparity in radar stereovision. The definition we adopt here is the closest to the one used in photogrammetry.

This form is more frequently used:

- either in particular cases simplifying the writing (for example, $B_z = 0$);
- or by approximations obtained by ignoring quantities h/H , B_z/H or sometimes B_x/x , or by limited developments.

We will discuss these particular cases below.

In [LEB 90] a solution is proposed that proceeds by equation linearization and iterative resolution with the help of successive approximations.

10.2.2.2. Simplified disparity/altitude relations

It does not appear clearly in equation [10.2] (because B_z is also present in the fraction denominator) that the expression is simplified when the basis is horizontal ($B_z = 0$); then:

$$h = \frac{2HB_x + 2Hp - \sqrt{4H^2 B_x^2 + p\Lambda}}{p + B_x} \quad [10.3]$$

with:

$$\Lambda = 8Bx(H^2 - x^2 + xB_x) + p(4B_x^2 + p^2 + 4pB_x) + 4p(H^2 - x^2 + xB_x)$$

This equation is simplified even more if we are placed within the plane wave front hypothesis (altitude of carrier H is significant compared to disparities p (altitudes h) under consideration); this approximation is valid in satellite imaging. We therefore obtain a particularly simple expression in accordance with angles θ' and θ'' :

$$h = \frac{p}{\cotan \theta'' - \cotan \theta'} \quad [10.4]$$

This formula is adopted in most works done with ERS, RADARSAT or Sir-B/Sir-C. It is valid for differences in altitude lower than 3,000 m approximately.

We can also give a limit development of expression [10.4]:

$$h = -2p \frac{x(x-B)}{BH} + 2p^2 \frac{3x^2 - 3xB + B^2}{HB^2} + O(p^3)$$

which points out the quasi-linear altitude dependence factor with disparity for significant altitude differences. This formula also shows the dependence of ratio h/p in relation to distance x : a constant disparity on the image does not correspond to a constant altitude on the ground².

10.2.3. Matching

10.2.3.1. Correlation

Following the example of photogrammetry, radargrammetry has generally adopted research concerning matching between image points of *correlation*.

For this, we define an often square neighborhood (window $v(i, j)$ with a size, for instance, of $(2n + 1) \times (2n + 1)$ pixels) around a position point (i, j) of reference image f . We seek the best correspondent on second image g by moving a similar window along the epipolar line (consequently, we will consider that this epipolar line is the distance line of the image, in accordance with the statements in section 10.2.1) and by evaluating various “similarity” measurements between these windows. The simplest measurement is the *correlation*:

$$\gamma_{fg}(k; i, j) = \sum_{l=-n}^n \sum_{m=-n}^n f(i+l, j+m) g(i+l+k, j+m). \quad [10.5]$$

This correlation is often preferred to the *centered correlation*, which is less sensitive to average level variations of the area under study:

$$c_{fg}(k; i, j) = \gamma_{(f-\bar{f})(g-\bar{g})}(k; i, j) \quad [10.6]$$

or better still, to the *standardized centered correlation*:

² This observation will also be made in Chapter 12 and it leads to the notion of orbital fringes (see section 12.3).

$$C_{fg}(k; i, j) = \frac{c_{fg}(k; i, j)}{\sigma_f \sigma_g} \quad [10.7]$$

where \bar{f}, \bar{g} and σ_f, σ_g measure respectively the averages and deviation-types of f and g on $\mathcal{V}(i, j)$.

The disparity p is then defined as:

$$p(i, j) = k_0 \text{ such that } G(k_0; i, j) > G(k; i, j) \forall k$$

for G equal to one of the three functions γ, c or C .

In general, only the sufficiently clear matching situations are chosen, namely those with strong measurement and which are clearly different from other candidates (for example, with measurement [10.7], we impose for $C_{fg}(k, i, j)$ not only to be a local maximum of C_{fg} , but also to surpass a certain threshold (0.5 or 0.75 for example). Similarly, it must be sufficiently higher than the other candidates, so that doubtful extremes may be avoided). The points that do not verify these conditions, and which would lead to weak or ambiguous measurements, are not matched and do not appear in the disparity chart or in the DTM.

In optical imaging, under the hypothesis in which both images f and g depend on each other by a linear relation (taking into account sensors and illumination variations) and under the hypothesis of additional noise, measurement [10.7] gives the best results on heterogenous scenes, while measurement [10.6], which is faster to calculate, is less satisfactory. These criteria can be interpreted as least squares criteria because, for example, the maximum of measurement [10.5] is equivalent to the distance minimum:

$$d_{fg}^2(i, j) = \min_k \left[\sum_{l=-n}^n \sum_{m=-n}^n (f(i+l, j+m) - g(i+l+k, j+m))^2 \right]$$

Taking this property into consideration, a measurement of type:

$$d_{fg}^1(i, j) = \min_k \left[\sum_{l=-n}^n \sum_{m=-n}^n |f(i+l, j+m) - g(i+l+k, j+m)| \right] \quad [10.8]$$

is often adopted for a quicker calculation. This is the property used in the first works on satellite radargrammetry starting with SIR-B [RAM 86].

In radar imaging, it is not possible to adopt the additive noise hypothesis. However, it is formula [10.7] that is most often used because of its simplicity. There are two ways of fighting against noise, which are very important in radar imaging:

- by increasing window size $2n + 1$, to the detriment of calculation time, mostly of measurement precision because the altitude obtained is the window altitude average; in optics, n is in the order of 2 or 3; in radar, it often has the value 5 or 10;
- by preceding the correlation action with an image-filtering step, which may reduce noise: in [SYL 97], better performances are measured with a Frost filter pre-filtering than without pre-filtering; better achievements are indicated than those obtained with filters considered more powerful. However, this step is not generally considered beneficial.

Nocera [NOC 96] proposed a better filter adapted to radar. In the case of multiplicative noise and signal in N looks following a probability law of \mathcal{X}^2 , the measurement is given by:

$$\gamma'_{fg}(k; i, j) = \sum_{l=-n}^n \sum_{m=-n}^n \frac{(2N-2)!N}{((N-1)!)^2} \frac{f(i+l, j+m)^{N-1} g(i+l+k, j+m)^{N-1}}{(f(i+l, j+m) + g(i+l, j+m))^{2N-1}} \quad [10.9]$$

where the multiplicative aspect of noise is taken into consideration by division with the estimated value of the signal (and not by subtracting). For a high number of looks (high N), Nocera estimates that this criterion is very close to that which gives equation [10.5]; in practice, he prefers it to the other criterion because it is easier to implement.

Hervet proposed to calculate the correlation of two windows whose statistics verify a Gamma law (see section 5.3.4.1) [HER 00]. With the ML criterion he obtains a complex expression, which does not admit an explicit solution. In a second

approach, he distinguishes between exact and fortuitous correlations by the presence of statistically dependent or independent pixels. Considering the hypothesis of Gamma distributions, he finds an expression similar to [10.7] in the case of independent law parameters, but he finds a still non-explicit and very difficult formula in the case of the same-parameter laws.

We still wish for a study, both theoretical and experimental, to define the criterion that is most suitable to a rapid and efficient radar image “correlation”.

10.2.3.2. Pyramidal approach

One way of improving the quality of altitude charts obtained by stereovision, while reducing calculation time, consists of working in a pyramidal manner. This technique, which is universally used in optical imaging, has also proven useful in radar imaging.

The principle is simple enough. From image $f(x, y)$, a pyramid of images $f^k(x, y)$ is created; this pyramid has decreasing resolution and is created by a low-pass filter of impulse response w_k .

$$f^k(x, y) = \iint f(x-t, y-u) w_k(t, u) dt du$$

This is achieved with a Gaussian filter, for example, which proves to be a good pyramidal tool:

$$w_k(x, y) = \frac{1}{2^k \sigma \sqrt{2\pi}} \exp\left(-\frac{x^2 + y^2}{2^{2k} \sigma^2}\right)$$

Then, we divide the image of a factor 2^k ; thus, the iteration of order k offers an image of the size $\frac{N}{2^k}$, with a maximal disparity of $\frac{P_{\max}}{2^k}$ if the original image is of size $N \times N$ and maximal disparity p_{\max} .

Then we start the matching process at level k (k takes values from 4 to 8). Since the image is strongly filtered, speckle effects are low and correlation is strong. Exhaustive research is conducted for each point on the entire range of possible disparities. An initial disparity chart P_k is obtained, giving a rough but reliable idea of the relief. We then move on to resolution $k - 1$. For each pixel, we now have an approximate disparity due to P_k at our disposal. This makes it possible to reduce the

research field of matchings. Thus, we build P_{k-1} . The process is continued until final resolution: $k = 0$.

Today this process is often used to obtain digital ground models by radargrammetry of very large scenes (for example, scenes of $8,000 \times 8,000$ pixels with RADARSAT).

10.2.4. Geometry effects connected with radar viewing

The importance of ground geometry in radar imaging has already been discussed. Indeed, ground orientation often has a more important effect than albedo in pixel radiometry. Three consequences derive from this aspect:

- matching actions performed in folding areas (those areas that appear in each image as local intensity maxima) will be incorrect (because these are not the same areas of the land creating folding) and they will need to be eliminated from the altitude chart (see Figure 10.4a);
- the same will happen with matching actions involving hidden parts and their borders (see Figure 10.4b);
- finally, even in the case of points that do not come from any of these categories, the radiometry variation could lead to matching errors.

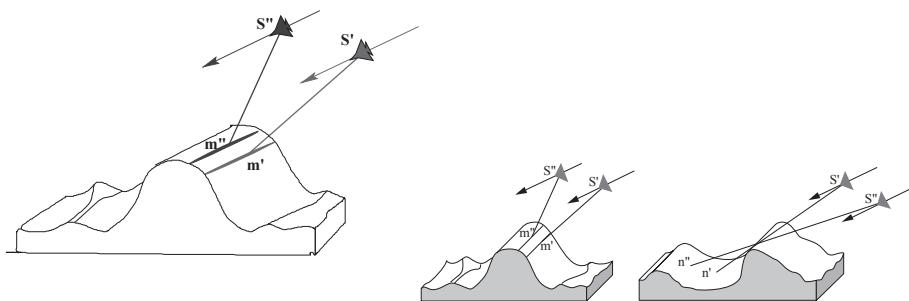


Figure 10.4. Left, folding effect: in each of the two images obtained by S' and S'' , a line will appear brightly, corresponding to a perpendicular incidence of the radar to the ground.

Certainly, these two lines will be matched by correlation techniques. However, they correspond to different points on the ground, m' and m'' and the altitude that will be deduced from the disparity obtained during matching will be incorrect.

Right, hidden parts effect: the 2 limits of shadows n' and n'' display very similar contrasts and will certainly be wrongly matched

The first two defects are particularly serious in the regions of strong reliefs and they may lead to altitude errors of several hundreds of meters. It will be difficult to avoid them with the present automatic correlation techniques, which seek similar pixel configurations in both images. In the absence of textures, the folding lines and shadows are the main candidates for the correspondence and it is very difficult to correct these errors. One solution consists of formerly detecting these areas (see Chapter 7 and section 13.2.3), then taking them away from the areas meant to be matched. Another solution is to take advantage of the geometric information brought by these shining lines and these shadows in order to deduce relative altitudes. This idea, which was presented early on by La Prade [LAP 69], does not seem to have been implemented yet in practice in a systematic way. We find traces of it in the works conducted in an interactive manner on buildings, for example.

The last defect is more subtle and apparently less serious, but it makes precision radargrammetry more complex. The orientation of surfaces leads to radiometry differences of the same point observed along 2 directions, in optical imaging as well as in radar imaging, because surfaces disperse different energies depending on their orientation³. This problem is more complex in radar imaging because the lighting and observation angles are sometimes modified at the same time between both observations. However, this problem is amplified even more by the “time-of-flight” effect, which directly connects pixel size (and quantity of energy received) to relative land orientation observed. Precision radargrammetry must take these effects into account. In the absence of good knowledge of bidirectional backscattering properties from observed scenes, it is difficult to compensate for the former effects. It is simpler to compensate for the latter effects, due to time of flight, by weighting radiometry results with surfaces involved in the correlation formula (formula [10.6], for example). Surfaces A_f and A_g are given by formula [11.5] and lead to a correlation expression with the following form:

$$\gamma''_{fg}(k; i, j) = \sum_{l=-n}^n \sum_{m=-n}^n \frac{(f(i+l, j+m) - \bar{f})(g(i+l, j+m) - \bar{g})}{A_f(i+l, j+m)} \frac{A_g(i+l, j+m)}{A_g(i+l, j+m)}. \quad [10.10]$$

However, A_f and A_g values, which depend on orientation, can only be calculated from knowing the DTM. This is precisely the objective of radargrammetry reconstruction. Therefore, it is convenient to work in an iterative manner and a first reconstruction is performed, for example, with equation [10.6]. It makes it possible to calculate orientation angles required for evaluating terms A_f and A_g . These evaluations make it possible to refine the DTM with the help of equation [10.10]; it is also possible for these operations to be iterated.

³ There is a single exception to this situation, namely that of Lambertian objects.

10.2.5. Filling hidden parts and non-matched areas

DTMs obtained with radargrammetry are often discontinuous; this happens both because of invalidated parts – for example, because of folding areas or hidden parts, as we have seen in section 10.2.4 – and because of weak correlations that often lead to rejecting doubtful matchings. In this case, the disparity chart or DTM must be filled in with the missing values. This is a delicate stage. If the missing points are not numerous and very isolated, an interpolation technique is often chosen, which “blocks” the holes and ensures a reasonable continuity. This cannot be applied to large hidden parts. We can use various procedures, for example:

- by advancing few hypotheses on the surface to be filled, it is possible to circulate known information to the unknown points. We proceed in an iterative manner by the following recurrence, supposing that all unknown altitudes have been arbitrarily set to value 0:

$$h^{k+1}(i, j) = h(i, j) \quad \text{if} \quad h(i, j) \neq 0$$

otherwise:

$$\begin{aligned} h^{k+1}(i, j) &= \alpha h^k(i, j) + (1 - \alpha) \bar{h}^k(i, j) \quad \text{if} \quad h^k(i, j) \neq 0 \\ &= h^k(i, j) \quad \text{otherwise} \end{aligned}$$

where $\bar{h}^k(i, j)$ is an average local estimator and α a parameter comprised between 0 and 1;

- with more information on landscapes geomorphology, we can introduce constraints that are more specific to a given area, for example, by favoring vertical planes (cliff edges, buildings), or horizontal terraces (paddy fields), or by confirming water flows in valleys, etc.

Despite these refining practices, the non-matched parts contribute considerably to a mediocre quality of DTMs obtained by radargrammetry and it is necessary to try to reduce them as much as possible, at the time of choosing the mission's parameters (particularly the incidence angles) as well as during image treatments.

10.3. Results of radargrammetry

Radargrammetry has considerably progressed in the past few years by making the most of radar imaging specificity and by delimiting itself from standard photogrammetry. The first convincing results of satellite radargrammetry are

attributable to [RAM 86] from SIR-B. However, the use of satellite radar RADARSAT, which proposes a variety of incidences, has made it possible to demonstrate the advantage of this type of process on a much larger number of landscapes because it is possible to choose a combination adapted to the incidence on certain landscapes (see Table 3.4). [MAR 97, DUP 98, SYL 98] put forward precise performance studies. We will indicate here a few conclusions that are valid today in relation to platforms such as RADARSAT, offering possibilities of variable incidence and resolution.

Opposite incidences are often considered too difficult to match and same-sense incidences are often preferred. However, two DTMs obtained, one with two rising incidences and the other with two descending incidences, can be used simultaneously in order to reconstitute a better-quality DTM, which is richer in reliable matchings and not marred with hidden parts. This strategy has been successfully used in [DUP 98].

In general, the lowest possible incidence views are preferred in order to reduce hidden parts [RAM 86]. For same side incidences, this leads to choosing the lower view of incidence that is as vertical as possible (for instance, the RADARSAT S1 mode). There remains then to choose the incidence of the strong viewing incidence. Altimetric precision is improved by enlarging the bases (therefore by using a strong incidence). The number of hidden parts and folding instances is increased; on the other hand, matching becomes more difficult. Therefore, the choice is made according to characteristics of the land to be measured: if it is mountainous or very low-textured, we choose a low stereoscopic angle in the order of 8 to 12° (which means that mode S4 or S5 of RADARSAT is preferred). On very flat land covered with crops or various vegetation, we benefit from choosing a stronger stereoscopic angle (from 15 to 25°, given for example by modes S6 or S7 of RADARSAT).

It is often considered that a DTM can be used when it leaves about 20% of non-matched pixels (this figure depends on many parameters of the application). In the case of stereovision on radar images, these performances are difficult to obtain and the reported figures often hover around 30%. This is one of the weakest points of radargrammetry.

On large areas, such as those covered by a RADARSAT scene, precisions obtained display good quality average values: biases are generally zero and deviation-types are in the order of 15 to 30 m for the 25 m spatial resolutions. However, extreme values may be strong, especially in the case of uneven terrain, where errors of several hundreds of meters may result from matching errors such as those we have mentioned in the above sections.

10.4. Conclusion

Radargrammetry has now earned the right to be mentioned among operational relief reconstruction techniques [SYL 97, SYL 98, ELI 00]. Its sources of error are not yet completely mastered but its mechanisms are well-understood and solutions have been found to several problems hindering its development. The use of radargrammetry in airborne imaging remains a delicate matter because of the necessity of rigorous carrier trajectory control. It is simplified in satellite imaging because trajectories are more constant and better known. The defects that mar DTMs obtained with radargrammetry still require the intervention of a human operator in order to correct possibly serious errors or in order to fill in non-matched areas. Resorting to other sources of information (DTM support, information obtained from radar-clinometry, SPOT imaging) may make it possible to avoid human intervention at times.

This page intentionally left blank

Chapter 11

Radarclinometry

Radarclinometry is the second technique used to build an elevation chart based on SAR imaging [TOU 00]. In contrast to the other two methods¹, radargrammetry (see Chapter 10) and interferometry (see Chapter 12), radarclinometry uses a single image. It is an easy method to implement, and working conditions, although restrictive, lead to reduced precision results. Radarclinometry can be considered a *qualitative* relief reconstruction method, insofar as the reconstructed reliefs are globally and morphologically similar to the real reliefs, but with elevations marred by large errors.

Radarclinometry is based on the *shape from shading* principle; this is a monoscopic image reconstruction method largely studied in computer vision domain [HOR 89]. It was first proposed for radar by Wildey, who applied it successfully to Venus landscapes captured by the Magellan space probe [WIL 86b, WIL 86a]. However, it has not been possible to validate these results by cartography in the field. While few researchers have been interested in this method since, we will mainly present the works of Frankot and Chellappa [FRA 87, FRA 90] and those of Paquerault [PAQ 97, PAQ 98] (see also [GUI 90, THO 91]).

Chapter written by Henri MAÎTRE.

¹ In section 11.5 a fourth method will be mentioned: radarpolarimetry, but this method is not well developed.

11.1. Radarclinometry equation

11.1.1. Shape from shading

The problems of shape from shading [HOR 89] are expressed by an equation with partial derivatives connecting reflectance R of an object to lighting function E by space derivatives: $p = \frac{\partial z}{\partial x}$ and $q = \frac{\partial z}{\partial y}$. Values R and E are known for each point of the object and elevation $z(x, y)$ is unknown for the problem that we seek to determine:

$$R(p, q) = E. \quad [11.1]$$

Actually, the problem most generally expressed by relation [11.1] is only rarely addressed. A more straightforward expression is preferred, corresponding to the case of cylindrical symmetry reflectance laws, for which:

$$R(p, q) = f(p^2 + q^2) = E$$

that leads to an expression called the *eikonal equation* which has enjoyed great attention in works about optics because it is directly attached to wave propagation. We have already encountered it at the beginning of this work (equation [1.9]). We write it here in the form:

$$\left(\frac{\partial z}{\partial x} \right)^2 + \left(\frac{\partial z}{\partial y} \right)^2 = g(x, y), \quad [11.2]$$

where $g(x, y)$ is a known function of x and of y .

In addition, with no loss in generality, the eikonal equation resolution for any reflectance distribution often comes down to a Lambertian reflectance, where the luminance is the same regardless of the surface orientation. It is this particular case therefore that has received the greatest attention in other works.

Unfortunately, this problem is ill-posed because in each point of the image, i.e. for each measurement performed, the user can only use a single measurement (measured radiometry) in order to determine the two unknowns represented by surface orientations (or both p and q derivatives). Because of this, there is an infinite number of solutions (generally they are conditions on continuity and derivability of the chosen surface, as well as limit conditions).

The solutions currently adopted in computer vision are: characteristic bands method [HOR 75], iterative methods [HOR 75], propagation methods [BRU 88] and viscosity solutions [LIO 91]. In practice, these methods make it possible to obtain satisfactory information on the study object's form, under certain conditions: on one hand, the signal should be less noisy; and on the other hand the object should be made of a single material, with the same reflectance in any point.

11.1.2. Shape from shading and radar

When shape from shading techniques are applied to radar images, four important differences are observed:

- the light source and receptor occupy the same position in radar imaging; this simplifies the equations by suppressing one of the angles, which is often unknown or less known;
- the signal is very noisy due to speckle;
- however, since the sensor is active (it is its own source), there are few parasite signals to disturb the measurement;
- radar imaging is much more sensitive than optical imaging to surfaces orientation and, it is much less sensitive to variations in the object's nature.

These properties lead us to select sounder, but probably less accurate, reconstruction techniques to handle radar images. After putting radarclinometry into equations, we will introduce the solution described in [PAQ 98] and results obtained on actual grounds.

11.1.3. Radarclinometry equation

Since the objective of radarclinometry is exploiting the dependence of radar imaging on observed surfaces orientation in order to determine ground slope and deducing, by integration, point elevation from the ground, it is important to connect the intensity of each pixel to the geometric land properties.

The radar wave is considered as a wave front plane to the Earth surface (approximation justified for satellite radar). A landmark will be chosen, represented in Figure 11.1, where axis Ox is vertical, and directions Ox and Oy are pointed along distances and azimuths. The radar is in plane Oxz and its incidence angle is θ . The *plot* is the parcel of land intercepted by the radar's resolution cell, generating the pixel. We presume it to be planar. Its orientation is given by two angles noted by α in accordance with distances and by β in accordance with azimuths.

The intensity sent by the plot in the direction of the sensor is a function of the plot size and orientation. These dimensions $L(\alpha)$ and $l(\beta)$ are expressed according to the resolutions in distance δ_d and in azimuth δ_a by:

$$L(\alpha) = \frac{\delta_d \sin \theta}{\sin(\theta - \alpha)} \quad [11.3]$$

$$l(\beta) = \frac{\delta_a}{\cos \beta} \quad [11.4]$$

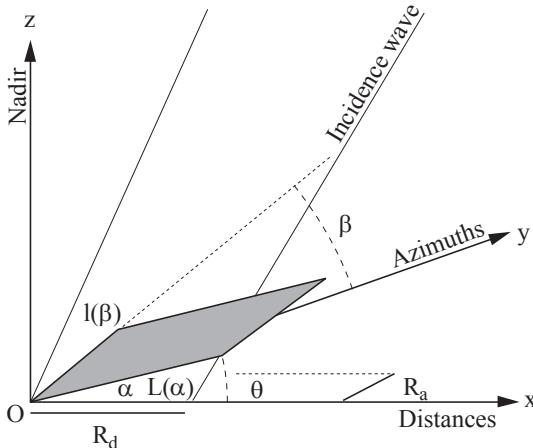


Figure 11.1. Representation of a plot, its dimensions and its two defining angles

The surface is then:

$$A(\alpha, \beta) = L(\alpha) \cdot l(\beta) = \frac{\delta_d \delta_a \sin \theta}{\sin(\theta - \alpha) \cos(\beta)} \quad [11.5]$$

The proportion of energy sent towards the sensor per surface unit is not the same, but it varies according to the nature of the soil. It is generally a function of lighting and observation angles (which are identical in the case of radar, where the transmitter is confused with the receptor). This point will be discussed in section 11.3.2, but in order to simplify the formalism, we will suppose that the land displays a Lambertian reflectance, which means that it depends only on the incidence angle, not on the observation angle. It is then possible to link intensity I received by a plot from the radar to its orientation parameters:

$$I(\alpha, \beta) = K \sigma^o \delta_d \delta_a \frac{\sin \theta \cos^2(\theta - \alpha)}{\sin(\theta - \alpha)} \cos \beta \quad [11.6]$$

where K is a constant regrouping all the parameters relating to radar calibration (see section 4.2) and σ^o , i.e. equivalent radar surface (see section 1.2.2.2), regroups the reflectance parameters related to the land. We will suppose $K\sigma^o$ to be known; this is also discussed in section 11.3.1. The objective of radarclinometry is to reverse equation [11.6] in order to determine values α and β , to determine the declivities associated with each plot and to integrate these declivities in order to determine plot elevations.

11.2. Resolution of the radarclinometry equation

11.2.1. *Wildey approach*

The first solution proposed is from Wildey [WIL 86b, WIL 86a]. It is based on a strong hypothesis, which makes it possible to suppress one of the two angular unknowns of equation [11.6]: Wildey presupposes that the land is locally cylindrical, or that there is only one non-zero curvature. From this hypothesis, he shows that the eikonal equation can be integrated line by line, from a boundary condition along one line. He proceeds with the direct integration of equation [11.2] point by point, by local estimation of slopes p and q from radiometry derivatives $I_x = \frac{\partial I}{\partial x}$ and $I_y = \frac{\partial I}{\partial y}$. A post-treatment step makes it possible to step away from the hypothesis of local cylindricality by calculating a second derivative corrective term on the profile thus obtained.

The Wildey method is relatively quick², but less robust; it is particularly sensitive to speckle noise and therefore it can only be applied to images strongly reduced by averaging. It has been tested on simulated images as well as on SEASAT images by using radiometry distributions from real data and adapted to ground coverage.

² Quick ... with current calculation methods, but in 1986 it was the long process time that seemed to Wildey as the main limitation for its use. Calculation time was approximately 20 hours for images of 630×630 pixels on a mini-computer.

11.2.2. Frankot and Chellappa approach

This is the most well-known method. It has generated abundant works and several experiments, in particular with SIR-B data [FRA 87, FRA 90]. It is based on supplementary data, knowledge of a digital low-resolution land model (i.e. with a resolution of 10 or 50 times lower than that attempted by the application).

This method researches, by considering the hypothesis of analyticity of z , i.e. the hypothesis that $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$, a solution³ minimizing:

$$\epsilon = \iint \left[I(x, y) - K \sigma^o \delta_d \delta_a \frac{\sin \theta \cos^2(\theta - \alpha(x, y))}{\sin(\theta - \alpha(x, y))} \cos \beta(x, y) \right]^2 + \lambda \left[\left(\frac{\partial p}{\partial x} \right)^2 + \left(\frac{\partial q}{\partial y} \right)^2 + 2 \left(\frac{\partial p}{\partial y} \right)^2 \right] dx dy \quad [11.7]$$

This method uses a regularization term based on the minimization of surface curves and it is weighted, according to the user's choice, by λ . The solution is found in an iterative manner.

Starting from a surface with very low resolution (in this case, a resolution given by raw DTM at our disposal), slopes p^k and q^k (the angles α and β) are therefore locally estimated for iteration k . From this estimation, a simulated radiometry $I^k(x, y)$ is deduced, with gradients calculated as follows:

$$I_x^k = \frac{\partial}{\partial p} I^k(x, y) \quad \text{and} \quad I_y^k = \frac{\partial}{\partial q} I^k(x, y).$$

making it possible to estimate slopes p^{k+1} and q^{k+1} in a more precise manner:

$$\begin{aligned} p^{k+1} &= p^k + \lambda [I(x, y) - I^k(x, y)] I_x^k \\ q^{k+1} &= q^k + \lambda [I(x, y) - I^k(x, y)] I_y^k \end{aligned}$$

³ In fact, instead of working on a terrestrial landmark (x, y, z) , Frankot and Chellappa remain within the radar mark captured by the viewing direction, azimuth axis and average in these two directions. The equations obtained are not very different from those above.

For each iteration, we make sure that surface $z(x, y)$ is regular, by an integrability constraint. For this purpose, surface $z(x, y)$ must locally fit with its tangent plane. This is obtained in the Fourier domain by imposing to the Fourier transform $Z(u, v)$ of $z(x, y)$ value:

$$Z(u, v) = \frac{K_x a_x^*(u, v) Z_x(u, v) + K_y a_y^*(u, v) Z_y(u, v)}{K_x |a_x^*(u, v)|^2 + K_y |a_y^*(u, v)|^2}$$

where Z_x and Z_y are local FT slope estimates, K_x and K_y are the chosen experimental weights to account for uncertainties of distance and azimuth measurements, and a_x , a_y are the distance and azimuth estimator factors (typically, $a_x = j \sin u$).

Frankot and Chellappa determine their problem's unknowns (parameters of equation [11.7]) in a systematic research by minimizing the difference between both members of equation [11.6].

The solution thus presented has interesting properties. The low frequencies of the resulting elevation chart are coarse DTM, which is at our disposal from the start. These are more reliable values than those obtained by radarclinometry because radarclinometry obtains low frequencies by integrating local slopes and therefore by accumulating uncertainties and errors. The integrability constraint ensures a regular form of the relief. The inversion, even if it is iterative, remains for an acceptable period of time because of a judicious use of rapid Fourier transforms. Nevertheless, the gradient descent method does not make it possible to find reliefs that are too different from the initial coarse relief because convergence is made towards a local minimum of equation [11.7]. It is therefore advisable to have a good initial estimation of the relief; radarclinometry enriches this estimation with high frequencies that were absent at initialization.

Frankot and Chellappa show good quality experimental verifications on SIR-B images.

11.2.3. Resolution of radargrammetry equation without a priori knowledge

The approach presented here belongs to Paquerault [PAQ 98]. It is one of the rare methods that do not use supporting data (supporting DTM, for example) but use high resolution radar signal. This approach is performed in two steps: the first gives a gross inversion of simplified equation [11.6], the second is a correction and regularization step.

11.2.3.1. Inversion of equation [11.6]

First, we assume that angle β is zero (it can be shown that, considering the usual values of radar orientation θ , this hypothesis has no great consequence on the resolution of equation [11.6]; this hypothesis is also made in [FRA 87] and mentioned in [WIL 86b]). In addition, the intensity of flat ground is supposed to be $I(0, 0)$. Then equation [11.6] becomes:

$$\mathcal{Q}(\theta, \alpha, \beta) = \frac{I(\alpha, \beta)}{I(0, 0)} = \frac{\sin \theta \cos^2(\theta - \alpha)}{\cos^2 \theta \sin(\theta - \alpha)} \quad [11.8]$$

It is inversed by taking as argument the only authorized binomial solution:

$$\alpha_s = \theta - \arcsin \left(-\frac{\mathcal{Q} \cos^2 \theta}{2 \sin \theta} + \frac{\sqrt{\mathcal{Q}^2 \cos^4 \theta + 4 \sin^2 \theta}}{2 \sin \theta} \right).$$

Pixel (i, j) is then responsible for a declivity $\delta h_{i,j}$ equal to:

$$\delta h_{i,j}(\alpha) = L(\alpha) \sin \alpha.$$

This declivity can be integrated in order to give the elevation at any point in the image. To simplify, if we choose to integrate based on a line along the increasing distances, elevation $H_{i,j}$ of a current pixel is expressed in relation to elevation $H_{i,0}$ of the first pixel of the line by:

$$H_{i,j}(\alpha) = H_{i,0} + \sum_{k=0}^j \delta h_{i,k}(\alpha). \quad [11.9]$$

However, since the radar image is marred by strong speckle noise, techniques less sensitive to clutter should be preferred to the integration along a simple line. Thus, in [PAQ 98] it has been proposed to apply several integrations along different orientation lines and to estimate the final value by the median of various values.

11.2.3.2. Regularization

At this point, a first digital model of a terrain is available, although the quality is still mediocre. We can then proceed to the second step of the restoration, i.e. a contextual Markovian reconstruction. This step has two objectives: first to

reintroduce angle β , which was ignored in the previous step, and second to correct the gross defects that are the result of weaknesses of this step and which mar the reconstruction. This will be carried out by solution regularization. At this stage, as in the case of all Markovian approaches, a form of energy is defined at each point of the image, energy which we will attempt to decrease by changing pixel elevations in an iterative manner. In addition, in the approach presented, this decrease will be made stochastically.

The energy is composed of two terms: the first is a term of connection to data, U_I , and the second a regularization term, U_v .

Connection to the given data is a penalty associated with the error between intensity I of pixel (i, j) and the intensity that would have a Lambertian orientation facet (α, β) . In order to take into account the strong granularity noise present in radar images, this term is bound by a difference higher than a γ threshold. For the iteration k , it takes the form:

$$\begin{aligned} U_I^{(k)}(i, j) &= [I(i, j) - K\sigma^o A(\theta, \alpha^k, \beta^k) \cos^2(\theta - \alpha^k) \cos^2 \beta^k]^2 \\ \text{if } & [I(i, j) - K\sigma^o A(\theta, \alpha^k, \beta^k) \cos^2(\theta - \alpha^k) \cos^2 \beta^k] \leq \gamma, \\ & = \gamma^2 \quad \text{otherwise.} \end{aligned}$$

The regularization term is directly related to point elevation and constrains the field to be continuous and horizontal:

$$U_I^{(k)}(i, j) = \sum_{i', j' \in \nu} \left(H^{(k)}(i, j) - H^{(k)}(i', j') \right)^2. \quad [11.10]$$

with \mathcal{V} neighborhood of (i, j) in 8-connectivity.

Therefore, the total energy has the value:

$$U^{(k)} = \sum_{i, j} \left[U_I^{(k)}(i, j) + \lambda U_v^{(k)}(i, j) \right].$$

Global energy on the entire image depends on elevations of each pixel. Its minimization cannot be done by traditional gradient descent techniques because of

its numerous local minima. We therefore adopt a stochastic optimization by simulated annealing⁴.

The initial elevation chart is supplied by the preceding stage (equation [11.9]). The algorithm, at iteration k , is as follows.

Starting from an elevation chart H^{k-1} and an annealing temperature T^{k-1} :

- we decrease temperature T (T^{k-1} becomes T^k);

- we randomly visit pixels. For each pixel:

- we propose an elevation modification $\delta H^{k-1}(i,j)$ uniformly drawn during an interval $[-\eta, \eta]$,

- we calculate angles α^k and β^k variations by the formulae (from Figure 11.1 brought down to ground geometry):

$$\alpha_{(i,j)}^k = \arctan \left(\left[\frac{\delta_d}{H_{(i,j)}^{k-1} - H_{(i,j-1)}^{k-1}} + \frac{1}{\tan \theta} \right]^{-1} \right), \quad [11.11]$$

$$\beta_{(i,j)}^k = \arctan \left(\frac{H_{(i,j)}^{k-1} - H_{(i-1,j)}^{k-1}}{\delta_a} \right), \quad [11.12]$$

- we deduce from this intensity $I(\theta, \alpha^k, \beta^k)$ by:

$$I^k(i,j) = K \sigma^o A(\theta, \alpha^k, \beta^k) \cos^2(\theta - \alpha^k) \cos^2 \beta^k,$$

- we calculate the contribution to energy variation of this pixel elevation change (i, j) by:

$$\delta U = U^{k-1} - U^k,$$

⁴ In simulated annealing, sites are visited consecutively in random order, and a modification of elevation is proposed for the site under examination. This modification is always accepted when it leads to an energy decrease. It may also be accepted even if it causes the energy to increase randomly, but the probability of being accepted decreases during iterations. The control of this probability is achieved by a parameter called *temperature*, which continues to decrease over time.

- this change is accepted if energy decreases. If energy increases, the change is accepted with a decreasing function probability of temperature T (according to the simulated annealing principle);

- we thus obtain an elevation chart H^k ;
- we increment k .

The algorithm stops when changes in the elevation chart become too low.

11.3. Determination of unknown parameters

One of the difficult problems of radarclinometry equation inversion is the determination of distributions and unknown parameters. In the case of a complex scene, we should determine these parameters locally. However, we will first suppose that the scene is homogenous, namely that parameters and distributions are unique for all points in the image, which is the hypothesis generally made in the published works. In most formulations, the determination of unknown parameters involves SER σ^o (or product Ka^o of equation [11.6]) and radiometry variation distributions in relation to the plot angle. The incidence angle at the center of scene θ is generally known exactly to satellite radars and its accurate expression for each pixel of the scene can be easily determined from the platform's resolution and elevation parameters. Consideration for the Earth's spherical shape in the determination of this angle is sometimes necessary for the strong incidences, but it is most often negligible considering the precision expected from radarclinometry techniques.

11.3.1. Determination of σ^o

The most conventional manner of determining σ^o is by learning from lands of known orientation. In order to have a reliable measurement, we should have a sufficient number of samples so that we can avoid speckle effects. This is often a problem because the average estimation with a precision of 5% often requires samples of more than 1,000 pixels.

Another way to proceed consists of conducting a statistical survey on the radiometry of the area under consideration, under the hypothesis, often verified, that the distribution of orientations is maximal for the horizontal and isotropic lands. Such an approach must take into consideration the bias introduced by radar geometry, which compresses distances oriented toward the radar and elongates those oriented in the opposite direction. Wildey [WIL 86a] proposes an accurate solution to this difficult problem of statistical estimation in the infrequent case in which slopes distribution is isotropic and Gaussian. Paquerault has shown experimentally

that the most commonly used robust estimators (average, mode, median) were not quite satisfactory [PAQ 98]; neither were those obtained on previously filtered images. She observes that a reliable estimation of σ^o amounts to causing the ground to tilt in the opposite direction to the radar, while a certain estimation has the reverse effect. She shows that for ERS, a 10% error on σ^o leads to a tilt of the ground of 1.8° . Therefore, she implements an iterative solution. From an accurate estimation of σ^o (given, for example, by the intensity distribution mode), we proceed to a reconstruction. Then we determine statistics on the relief obtained. For example, on mostly plane ground, elevation average and variance will be minimal for the correct σ^o value. This value is obtained by a decrease of the standard gradient.

11.3.2. Determination of backscattering distribution

The backscattering distribution expresses σ^o variations according to, on the one hand, angles created by the source, and by the receptor on the other hand, with facet average. In the radar case, these angles generally overlap. In all the preceding calculations, we have adopted a Lambertian backscattering distribution. Lambert's distribution expresses σ^o variations in relation to the only angle of the source (from now on we will designate it by w) and therefore independently of the observation direction. It also expresses isotropy of the phenomenon, in relation to both the incidence and detection angles:

$$\sigma^o(\omega) = \sigma^o(0) \sin^2 \omega \quad 0 \leq \omega \leq \frac{\pi}{2} \quad [11.13]$$

There are different opinions on the capacity of this distribution to reflect the Earth's surface properties: for some, this is a simple and good approximation, and for others a second best. Let us recall, however, that since the backscattering distribution also depends on wavelength, on polarization, on dielectric nature and on its roughness, it is quite difficult to find a generally satisfactory distribution. Inversely, it is often possible that different empirical laws should be well adapted to the same experimental data.

Clapp, in 1946 (cited in [LON 83] p. 75), was the first to propose a Lambertian model, which is better adapted to his experimental data:

$$\sigma^o(\omega) = \sigma^o(0) \sin \omega \quad 10^\circ \leq \omega \leq 80^\circ$$

This is called the second Clapp model, but he also proposed model $\sigma^o(w) = cste$ (third Clapp model), following a mathematical modeling by joint

spheres. These various models are adapted to various types of vegetations. They have been extended and justified theoretically by Keydell [KEY 82], who proposes a formulation generated by parts:

$$\sigma^o(\omega) = \sigma^o(0) \frac{\sin^{\nu+1} \omega}{\cos^\mu \omega} \quad \text{if } 0 \leq \omega \leq \omega_g$$

$$\sigma^o(\omega) = \sigma^o(0) \frac{\sin^{\nu+1} \omega_g}{\cos^\mu \omega_g} \quad \text{if } \omega_g \leq \omega \leq 90^\circ$$

11.4. Results

We present a few radarclinometry results, obtained with the Paquerault technique on Guyana landscapes, which are quite favorable to the use of this technique because they present a moderate relief and a regular and homogenous plant covering: a dense equatorial forest (Figure 11.2). In general, the Amazonian forest is a good example of the use of radarclinometry. Covered with clouds a large part of the year, with rather difficult land access, it is badly mapped and somewhat inaccessible to standard means of altimetric elevation.

Several observations on these results can be made:

- the reliefs are reasonably well-rendered: the reliefs of the small mountain massif are finally detailed and well reproduced, irregularities in the low plateau are also detected, high points and valley lines are well placed, slopes are correctly oriented, etc.;
- the first stage (Figure 11.2b) is unusable, in fact, because of its strong lineage, but it displays finer details than the final image (Figure 11.2c), which is highly filtered. Actually, this is the price to be paid in order to eliminate the lineage;
- a quantitative study of the final elevation chart properties shows that altimetry precision is mediocre. Elevation deviation types are of the order of 30 to 50 m for the total change in elevation of approximately 400 m [PAQ 97].

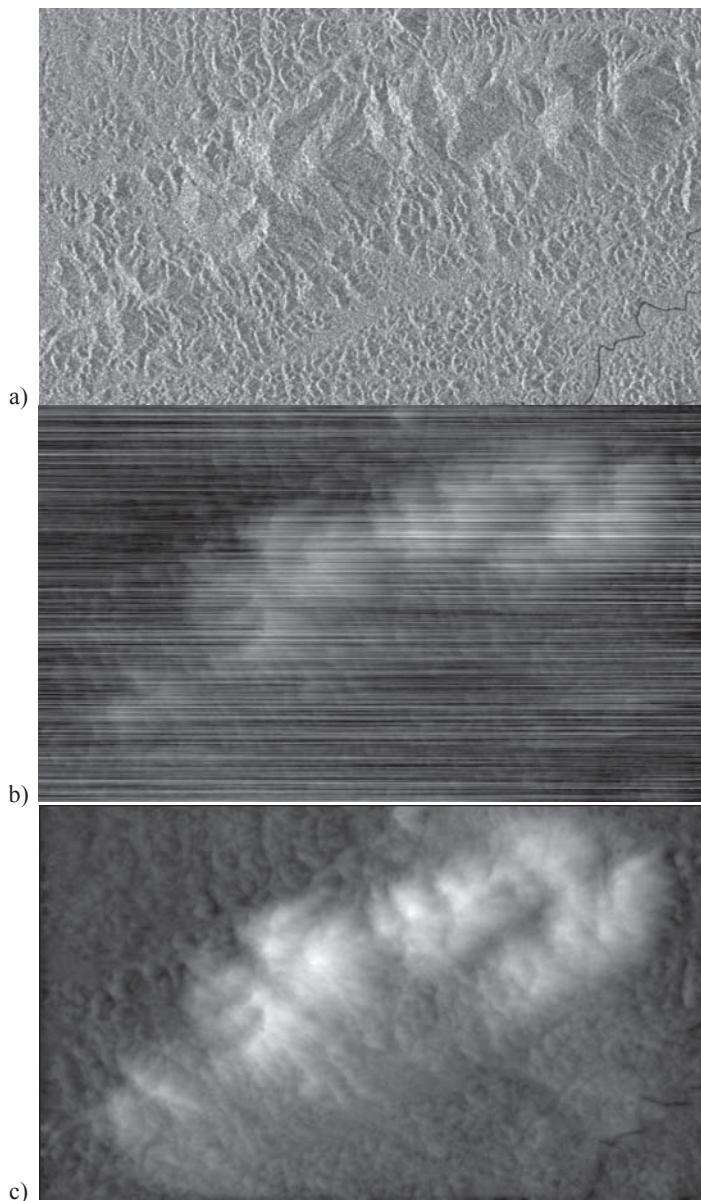


Figure 11.2. Excerpt from an ERS scene, with a 25 m resolution; Plomb mountains in Guyana (© ESA). Center: elevation chart initialization by lines integration (displaying a strong lineage); below: final elevation chart after Markovian restoration (processing: S. Paquerault)

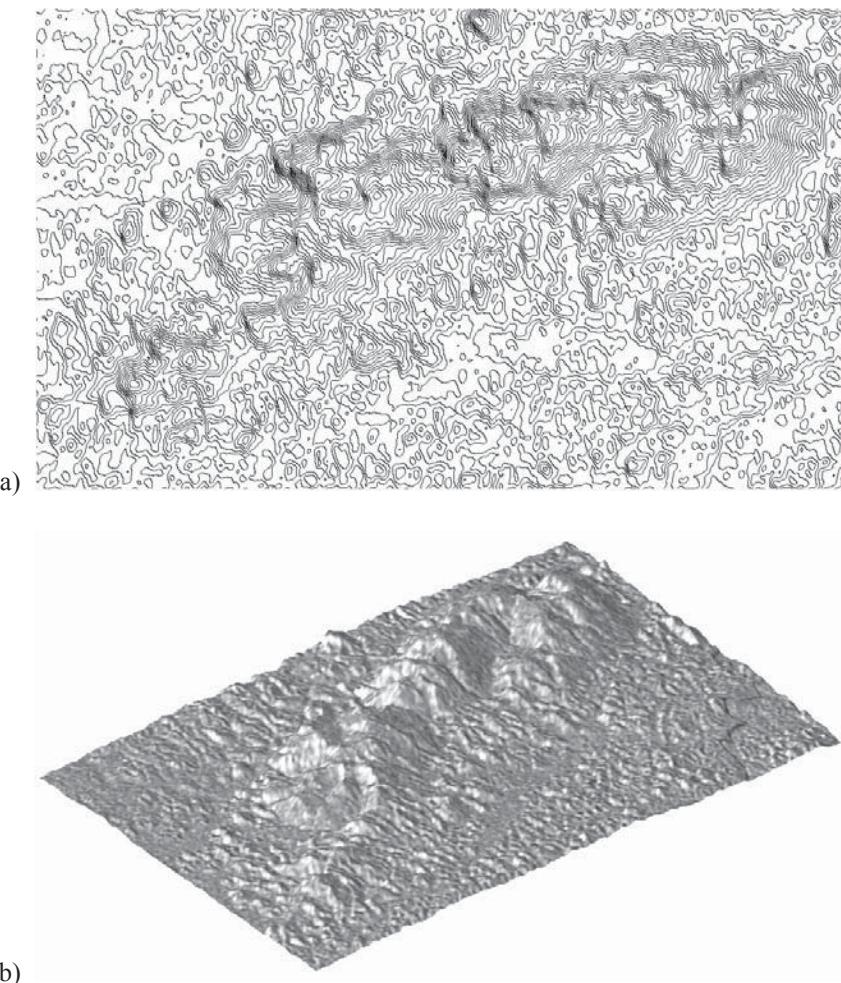


Figure 11.3. Level curves obtained from Plomb mountains from the elevation chart calculated by radarclinometry and perspective reconstruction of the ground (processing: S. Paquerault)

From these observations we can deduce that, in the current state of the art, radarclinometry is an efficient technique used for reconstituting general landscape morphology, but that it does not make it possible to obtain a quantitatively reliable altimetry.

reference to Figures 11.4 and 11.5	satellite	band	wavelength	incidence angle	polarization
a	ERS-1	C	5.66 cm	23°	VV
b	JERS-1	L	23.53 cm	32°	HH
c				25°	
d	RADARSAT	C	5.66 cm	32°	HH
e				45°	

Table 11.1. Image parameters used for the comparison of Figures 11.4 and 11.5

We discuss other properties of radarclinometry in the images in Figure 11.4. These images consist of 5 views obtained under different conditions: various sensors, under different angles and in various wavelengths. Table 11.1 indicates the parameters used for building these images. They represent a sub-area of Figure 11.2. Reconstructions are displayed in Figure 11.5. First, it is observed that there is a good concurrence of reliefs thus reconstructed, but there are also differences, which may be analyzed as follows:

- when only incidence angles differ (this is the case for the series of images obtained from RADARSAT), it is observed that the reconstruction is better for weak incidences. This happens because, if relief effects are less marked in the image-sources, the role of speckle is considerably weaker;
- the role of wavelength is not so marked but the comparison between RADARSAT for $\lambda = 5.66$ cm and 32° and JERS-1 for $\lambda = 23.53$ cm and 36° shows that longer wavelengths give good results under a dense plant cover and that the penetration difference of the cover is quite capable of counterbalancing the effects of a less favorable angle;
- the experiment led with these sensors does not make it possible for us to conclude on the advantage of a preferable polarization mode.

11.5. Radarpolarimetry

In principle, radarpolarimetry relies on different physical properties compared with radarclinometry, but it is very similar to radarclinometry in its implementation. As indicated by its name, this method uses multi-polarization data and it makes use of the different reflection properties of various polarizations of a wave that is incident on a plane surface in relation to its incidence. These properties have been discussed in section 1.3 and illustrated in Figure 1.2. Principles of the method were presented in [SCH 96].

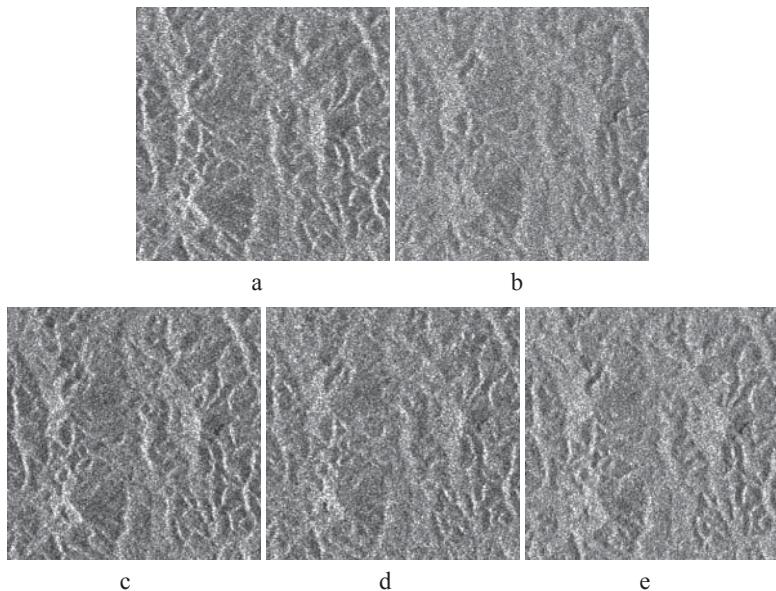


Figure 11.4. Five images of the same massif in Guyana obtained by various satellites: in a by ERS-1 (© ESA), in b by JERS-1 (© NASDA), in c, d and e by RADARSAT under incidences of 25° , 32° and 45° respectively (© ADRO)

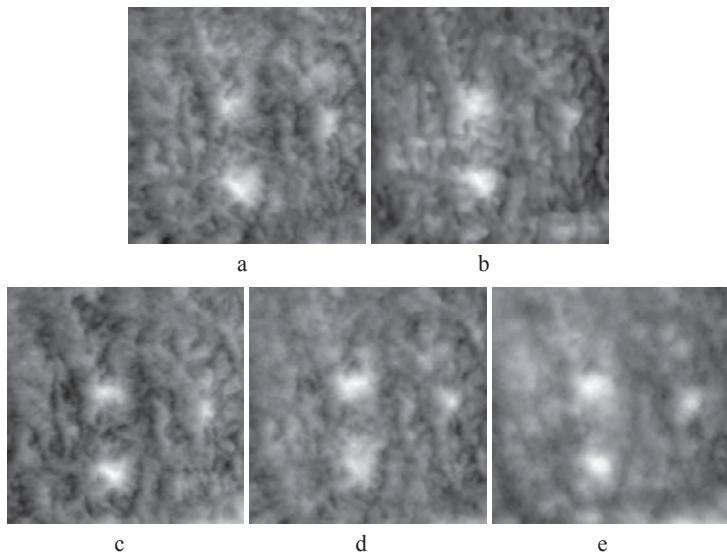


Figure 11.5. DTM obtained from the five images of Figure 11.4

When we examine the 4 normalized equivalent surface terms σ° of a plane and rough material, we observe that ellipticity polarization angle χ (see Figure 1.3) varies in an almost linear manner with angle β , formed by the plot in the azimuth direction (see Figure 11.1) as well as with angle α . As in the case of radarcinometry, each pixel only supplies one measure of χ in order to determine unknowns α and β . However, at the expense of certain regularity hypotheses on the surface (similar to those used in the eikonal equation inversion), it is possible to go back to the facet orientation and to the point's elevation. Schuler proposes a gradient technique able to reconstruct the surface from a flat ground hypothesis or from a rough ground digital model. In order to render the inversion more constant, image pixels are strongly averaged leading to a significant resolution degradation.

Further results are presented in [SCH 98], particularly by using two images acquired in cross directions. It should be observed that this technique, as in the case of radarcinometry, requires homogenous surfaces and seems particularly suitable to bare and dry grounds (as opposed to radarcinometry, which is more efficient on densely covered areas). It has been used to compensate the relief effects in SAR analysis [SCH 00]. It has also been used by the same authors in order to characterize the state of the sea [AIN 00].

Based on our knowledge, no other work has been dedicated to this technique.

11.6. Conclusion

Radarclinometry, just as with radarpolarimetry, has not yet been proven to provide precision altimetry by itself. It makes it possible, in favorable coverage conditions, to give qualitatively satisfactory information at very low expense. It may also be a very useful complement to other relief reconstruction techniques, in situations in which they do not work very well: forest coverage for interferometry, savannah or steppe bare lands for radargrammetry. The use of several images by radargrammetry and radarcinometry simultaneously is also proposed in [THO 91] and we show that it is possible to considerably improve radargrammetry performance alone.

Chapter 12

Interferometry

Although it is relatively recent, radar interferometry technique is already widely used by the entire radar remote sensing community. The principle was described by Graham in 1974 and validated on SEASAT and SIR-B data. However, it is only since the launching of the first European radar satellite ERS-1 in 1991 that it has rapidly developed. Since then, data acquisition by tandem ERS 1 & 2 (two images obtained by two sensors during a one-day interval) and, more recently, by the SRTM mission (two antennae placed on the shuttle) have enabled the implementation of this technique on a planetary scale.

Interferometry uses phase information contained in radar images as a very sensitive means of distance measurement, or more exactly of distance variations, because it makes sense only in a differential manner. Finite analysis of phase difference between two radar images makes it possible to reconstitute land forms or to measure land movements that may have taken place between several data collections with millimetric precision.

Currently, many laboratories use interferometry (or simply phase coherence, which is its sub-product) for their applications; these applications themselves are numerous in geophysics, topography, geology, glaciology, hydrology, volcanology, forestry, gas or oil exploitation, or the management of mineral resources in general.

12.1. Interferometry principle

12.1.1. Phase and interference

Each pixel of a complex radar image involves two sets of information: radiometry (wave amplitude) and phase.

Radiometry represents ground reflectivity to microwave: this value constitutes *stricto sensu* the image. Phase is the sum of two terms:

$$\varphi = \varphi_{specific} + \varphi_D$$

– the phase term that accounts for the interaction between wave and land, and which is called the specific pixel phase and is noted by $\varphi_{specific}$; it is connected to the physical nature of the land and to the look angle (see Chapter 1);

– the phase term that accounts for the round trip propagation; each time the wave crosses a distance equal to wavelength λ , its phase increases with $2\pi : \varphi_D = \frac{4\pi}{\lambda} |\vec{D}|$, where \vec{D} conventionally represents the range from the source to the center of the pixel.

The specific pixel phase is deterministic (therefore it is *a priori* possible to reproduce it), but it corresponds to a misknown variable, which for this reason can be seen as random, and which is at the origin of speckle (see section 5.2.1).

Therefore, we cannot use the phase as geometric information if we only have one image at our disposal. On the other hand, if we calculate the phase difference between the two images obtained in the same geometric conditions (same look angle, same speckle) and, if we can assume that ground nature has not been changed between the two data acquisitions, we can eliminate the specific pixel phase term. This then remains a purely geometric information on the range difference between both data acquisitions. By indicating by 1 and 2 the signals issued from both images, we obtain a phase difference:

$$\begin{aligned} \delta\varphi &= \varphi_2 - \varphi_1 = \left(\varphi_{specific} + \varphi_{|\vec{D}_2|} \right) - \left(\varphi_{specific} + \varphi_{|\vec{D}_1|} \right) \\ &= \frac{4\pi}{\lambda} \left(|\vec{D}_2| - |\vec{D}_1| \right) \end{aligned} \quad [12.1]$$

This measure is very accurate (of the order of a wavelength fraction or at centimeter level), but is only known modulo the wavelength (module λ). Therefore, the image of these phase differences will be made up of fringes. This image is called an interferogram. Reaching the information on range difference from the interferogram is achieved by fringe development, an often delicate operation which will be examined in Chapter 13.

12.1.2. Construction of an interferometric product

The interferometric product construction (i.e., interferogram + coherence image pairing) is the result of a succession of stages that can be described schematically in the following manner:

– *local image registration*. Its objective is to accurately superimpose one pixel (x, y) of image 1 to the corresponding pixel in image 2. In order to achieve this, the geometric transformation to apply to image 2 is to be determined so that, in an ideal case, all the elementary reflectors contributing to specific phase $\varphi_{specific}(x, y)$ during the formation of image 1 should be found exactly in the same pixel (x, y) of image 2. If this transformation is found, then it will be possible to exactly apply equation [12.1] in order to obtain the phase difference on a given point (x, y) .

This geometric transformation is first obtained by an analysis of acquisition conditions, namely radar tracks. It will be developed in section 12.3. In this regard, bi-static one pass systems (which carry both reception antennae on the same platform) are simpler to implement than the preceding two passes systems. Fine transformation parameters (translations usually¹), which give the best superposition, are then experimentally researched. This can be performed by correlation techniques, by spectral maximization, or by fringe observation. The expected precision of image registration is significantly lower than a pixel (between a tenth and a hundredth of a pixel) [GAB 88, FOR 95]. This requires a finite over-sampling of the signal. At this level, it is indispensable to take into account the exact processes of image construction (see Chapter 2);

– *global image registration*. The image registration from the previous step determines a geometric transformation (for example, a translation of parameters: δx and δy), which is valid in the neighborhood of a pixel (x, y) . If it had been repeated across the entire image, slightly different transformations would have been found in the pixels dispersed on the scene to be analyzed. Since we wish to obtain a sub-pixel image registration in all points of the scene, we should determine a global transformation which, applied to the complete scene, reproduces the preceding result. This is generally obtained by multiplying the previous local image

¹ Note that the experiments on SIR-B data dealt with data acquired on non-parallel tracks.

registrations in various points of the scene, and then interpolating all the points of the field thus found by an elastic transformation (for example, a transformation in which translation parameters δx and δy are low-degree polynomials in x and y);

– *interferogram construction*: the interferogram is calculated by finding term $\delta\varphi(x, y)$ from equation [12.1] at any point of the scene. In practice, from two complex signals $z_1(x, y)$ and $z_2(x, y)$, we calculate cross-correlation:

$$\rho(x, y) = \frac{z_1(x, y) z_2^*(x, y)}{\sqrt{|z_1(x, y)|^2 |z_2(x, y)|^2}} \quad [12.2]$$

If calculated in a single pixel, this value always has module 1. In order to estimate the correlation amplitude in a similar way, it is necessary to perform a local estimation (its exact writing will be seen in equation [12.4]). It gives us an expression which is written:

$$\rho(x, y) = d(x, y) e^{j\delta\varphi(x, y)} \quad [12.3]$$

The phase term $\delta\varphi(x, y)$ carries the desired information. The phase, measured between $-\pi$ and π is generally quantified on 8 bits and the interferogram, chart of $\delta\varphi(x, y)$ for any point (x, y) , appears then as the right image in Figure 12.3, namely as an image of wrapped phase;

– *coherence image construction*: we want that the coherence image to reflect in each point (x, y) its capability to supply comparable phases. All sources of uncertainty affecting the wave phase must then intervene in this term. The term $d(x, y)$ of expression [12.3] is a good coherence measure. In general, it is this relation that means coherence, but there are also other expressions, sometimes more empirical, which are obtained by means of a mere comparison of adjacent pixels phases.

12.1.3. What type of information does the interferogram carry?

The interferogram contains a certain quantity of information at the basis of interferometry applications; these applications will be discussed in section 12.4. We mention them here:

– information on elevation: reliefs create differences between the two images if orbits are not identical (stereo effect). These differences will be interpreted as a

phase difference of one image in relation to the other, which is proportional to land elevation. Then, interferogram fringes can be considered as a topographic map;

- information on land movement: the phase is proportional to the range. If the land is altered, it swells or it sinks between two radar acquisitions (earthquake, collapse because of groundwater depletion, landslides, etc.), modifications will be seen directly in the interferogram, with centimeter precision. Therefore, the interferogram can be interpreted as the image of land movements, where fringes are *iso-movement* lines;

- information on coherence: when interferometry hypotheses are not verified, i.e. when the physical or geometric nature of the land has changed (liquid surface, vegetation area, look angle variation), the pixels' phase differs from one image acquisition to another (various speckles) and the interferogram is locally noisy. Measurement of phase stability, called *coherence*, is a very finely tuned indicator of land changes, even if slight.

12.2. Interferogram modeling

12.2.1. Statistical laws of interferograms

Images used for assembling an interferogram are polluted by a strong speckle noise. This is the result of numerous complex factors, as we have seen in the previous chapters. Some of these factors change from one image to another and introduce a difference in the speckle disturbing phase difference estimation $\delta\phi$ and coherence d of equation [12.3].

In order to evaluate these terms in a more reliable manner, it is natural to estimate equation [12.2] no longer on one pixel, but on a small window of L neighboring pixels. Can we estimate the gain in precision which we have a right to expect from this operation?

By using the methods developed in Chapter 5, we can consider each image as a channel of signal $\mathbf{Z}(x, y)$ and the interferometry pair as a vector. We are then placed in the particular case of a circular Gaussian model with $N = 2$ channels, which has been analyzed in a multi-look framework in section 5.3.5. For L samples, we observe a signal $\mathbf{Z}_k(x, y)$ ($k \in 1, \dots, L$). It is possible to construct empirical covariance matrix Σ_z :

$$\Sigma_z = \frac{1}{L} \sum_{k=1}^L \mathbf{Z}_k(x, y)^t \mathbf{Z}_k^*(x, y) = \begin{pmatrix} I_1 & I_{12}e^{j\delta\varphi} \\ I_{12}e^{-j\delta\varphi} & I_2 \end{pmatrix}.$$

and then empirical coherence is expressed by the relation:

$$\hat{p} = de^{j\delta\varphi} = \frac{\sum_{k=1}^L z_{1,k} z_{2,k}^*}{\sqrt{\sum_{k=1}^L z_{1,k} z_{1,k}^*} \sqrt{\sum_{k=1}^L z_{2,k} z_{2,k}^*}} = \frac{I_{12}}{\sqrt{I_1 I_2}} e^{j\delta\varphi} L \geq 2. \quad [12.4]$$

where d expresses the empirical coherence degree and $\delta\varphi$ the empirical phase difference. These formulae extend the expressions of d and $\delta\varphi$ given in [12.2] and [12.3] to the case of estimation on L samples.

Their distribution laws are written as follows:

– *distribution of empirical phase difference* (according to formula [5.12]):

$$\begin{aligned} p(\delta\varphi | D, L) &= \frac{(1-D^2)L}{2\pi} \left({}_2F_1 \left(1, L; \frac{1}{2}; D^2 \cos^2(\delta\varphi) \right) D \cos(\delta\varphi) \right. \\ &\quad \left. + {}_1F_0 \left(L + \frac{1}{2}; -D^2 \cos^2(\delta\varphi) \right) \frac{\Gamma(\frac{1}{2}) \Gamma(L + \frac{1}{2})}{\Gamma(L)} \right) \\ &= \frac{(1-D^2)L}{2\pi} \frac{1}{2L+1} \cdot {}_2F_1 \left(2, 2L; L + \frac{3}{2}; \frac{1+D \cos(\delta\varphi)}{2} \right) \end{aligned} \quad [12.5]$$

– *distribution of empirical coherence degree d* (formula [5.13]):

$$p(d | D, L) = 2(L-1)(1-D^2)^L d(1-d^2)^{L-2} {}_2F_1 \left(L, L; 1; d^2 D^2 \right)$$

Illustrations of these laws were given in Figures 5.15 and 5.16.

12.2.2. Estimation of empirical coherence

Estimation of parameters d and $\delta\varphi$ of empirical coherence is carried out in practice by averaging magnitudes on a neighborhood of each pixel (equation [12.4]) \mathcal{V}_{xy} :

$$d(x, y) = \frac{\left| \sum_{x', y' \in v_{xy}} z_1(x', y') z_2^*(x', y') \right|}{\sqrt{\sum_{x', y' \in v_{xy}} |z_1(x', y')|^2} \sqrt{\sum_{x', y' \in v_{xy}} |z_2(x', y')|^2}}$$

$$\delta\varphi = \text{Arg} \left(\sum_{x', y' \in v_{xy}} z_1(x', y') z_2^*(x', y') \right)$$

The choice of a neighborhood is one of the most delicate points of interferometry techniques. In practice, it is necessary to take into account the following elements:

- data sampling: in general, the range step is larger than the azimuth step;
- degree of coherence between channels (if this degree is low, a large window is required, as will be seen in Figure 12.10);
- homogeneity of the scene (a low dimension of the window makes this condition less critical).

12.3. Geometric analysis of data

We have seen that phase difference $\delta\varphi$ (equation [12.1]) is directly linked to the range difference between both acquisitions. The purpose of this section is to connect this phase difference to sensor data (relative positions of the two tracks in the case of a traditional satellite, geometric characteristics of the double antenna in the case of SRTM mission) and to land data (local elevation).

Let a reference plane (Oxz)², which is orthogonal to satellite tracks (assumed to be parallel in an initial approximation) and a reference ground point $A = \begin{pmatrix} X \\ 0 \end{pmatrix}$.

The two acquisitions correspond to the following tracks:

- the master satellite S_1 track, located in $\begin{pmatrix} 0 \\ H \end{pmatrix}$, in such a way that look angle θ is defined by $\tan \theta = \frac{X}{H}$;

² The calculation we complete here has already been carried out, under different hypotheses, in the chapter on radargrammetry (see section 10.2.2), leading to equation [10.1].

– the slave satellite S_2 track, located at a distance of $\begin{pmatrix} B_x \\ B_z \end{pmatrix}$ from satellite S_1 . This distance is called interferometric baseline.

Let a second ground point : $B = \begin{pmatrix} X+x \\ z \end{pmatrix}$.

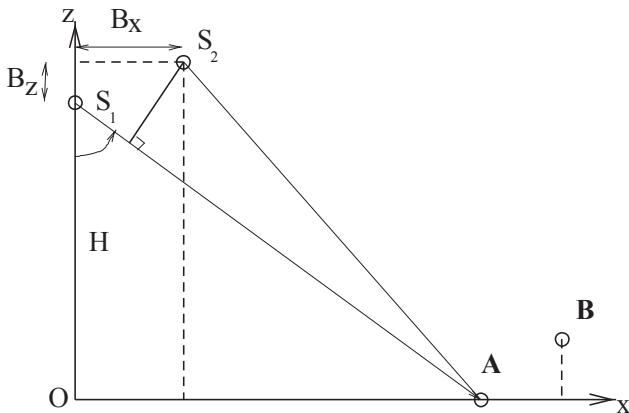


Figure 12.1. Interferometry point of reference. Tracks of satellites S_1 and S_2 are perpendicular to the figure

Now calculate difference δR between differences of ranges $D_{S_1A} - D_{S_2A}$ and $D_{S_1B} - D_{S_2B}$. Using the expressions:

$$D_{S_1A} = \sqrt{X^2 + H^2}$$

$$D_{S_2A} = \sqrt{(X - B_x)^2 + (H + B_z)^2}$$

$$D_{S_1B} = \sqrt{(X + x)^2 + (H - z)^2}$$

$$D_{S_2B} = \sqrt{(X + x - B_x)^2 + (H + B_z - z)^2},$$

by a second order expansion [ABD 98], we obtain δR approximation:

$$\begin{aligned}\delta R &= \frac{B_x \cos \theta + B_z \sin \theta}{\sqrt{X^2 + H^2}} (x \cos \theta + z \sin \theta) \\ &= \frac{B_{orth}}{\sqrt{X^2 + H^2}} (x \cos \theta + z \sin \theta)\end{aligned}\quad [12.6]$$

where the orthogonal component of the baseline in relation to the looking direction has been noted by B_{orth} :

$$B_{orth} = B_x \cos \theta + B_z \sin \theta$$

Two terms are observed in expression [12.6]:

– $\frac{B_{orth}}{\sqrt{X^2 + H^2}} x \cos \theta$: this term depends only on the positions of both satellites and

on the viewed point, while the Earth is assumed to be flat. It is at the origin of orbital fringes, also called flat Earth fringes, equivalent to air wedges fringes in optics. These fringes are parallel to the track. Modulation by orbital fringes may be similar to a device function;

– $\frac{B_{orth}}{\sqrt{X^2 + H^2}} z \sin \theta$: this term depends on the viewed point elevation. It is at the

origin of topographic fringes and therefore it is that which enables a determination of the scene's reliefs.

In the case where $b = B_x$ and $B_z = 0$, which is a rather traditional situation in satellite imaging, we find a better-known similar relation, that by Lin:

$$\delta R = \frac{b}{\sqrt{X^2 + H^2}} x + \frac{XHb}{\left(\sqrt{X^2 + H^2}\right)^3} z$$

If phase difference $\delta\varphi$ for both points A and B is known, we deduce δR , and elevation z , using equation [12.1], assumed that orbital parameters H , X , B_x , and B_z are precisely known.

Since the phase is known at only about 2π , we see that there is a specific elevation so that, on a wave front (considered as a plane), the phase rotation is equal

to 2π and therefore the path difference (equation [12.1]) is equal to $\frac{\lambda}{2}$. This elevation, noted by z_{amb} , is called ambiguity height.

Knowing that the wave front points verify $z \cos\theta - x \sin\theta$, by applying relation [12.6], we obtain:

$$\frac{B_x \cos\theta + B_z \sin\theta}{\sqrt{X^2 + H^2}} z_{amb} = \frac{\lambda}{2},$$

or

$$z_{amb} = \frac{\lambda X}{2B_{ortho}} = \frac{\lambda |\vec{D}| \sin\theta}{2B_{ortho}}.$$

Insofar as the elevations are low compared to the ambiguity height and in order to enable an accurate analysis of reliefs, it is tempting to increase the baseline. However, a new constraint appears if we want to ensure that both pixels remain in phase between both viewings. In fact, within a resolution cell (of dimensions δ_d in distance and δ_a in azimuth), ranges for the nearest point to the sensor and for the most distant one differ. If this difference of ranges differs up to λ between the two acquisitions, the speckle between them will be totally decorrelated, yielding a lack of coherence. On flat land we introduce the notion of *critical baseline*, of components $\begin{pmatrix} B_{Cx} \\ B_{Cz} \end{pmatrix}$, from the relation:

$$\frac{|B_{Cx} \cos(\theta) + B_{Cz} \sin(\theta)| \delta_d \cos(\theta)}{\sqrt{X^2 + H^2}} = \frac{\lambda}{2}. \quad [12.7]$$

12.4. Applications of interferometry

12.4.1. Elevation measurement

In traditional stereovision, two images are acquired with a strong look angle difference in such a manner that geometric alterations from one image to another are produced. These alterations depend on the ratio between baseline and height (noted by B/H, which is chosen close to 1; see section 10.1.2). They are proportional to

reliefs. In the case of interferometry, baselines are very low (B/H is in the order of 10^{-4}), geometric alterations are barely measurable, but they are enough to make the phase turn. One half-wavelength geometric distortion between the two images produces a phase turn. Therefore, it is no longer a shift, measured by correlation, which gives the elevation, but a phase difference. The first advantage is that the images are perfectly identical and therefore easily superimposed (no setting in epipolar geometry, no or few disparities, etc.). The second advantage is that the phase is measurable on the totality of the image, while in traditional stereovision correlation is quite ambiguous sometimes, mainly in uniform areas. On the other hand, the drawback is that this measurement is known only in module 2π , which creates fringes that should be unwrapped.

This is illustrated in Figure 12.2, where points A and B, at the same range from satellite S_1 , although they are not at the same altitude, display a step difference δl for satellite S_2 . This difference, if it is equal to a half-wavelength, creates a dark fringe.

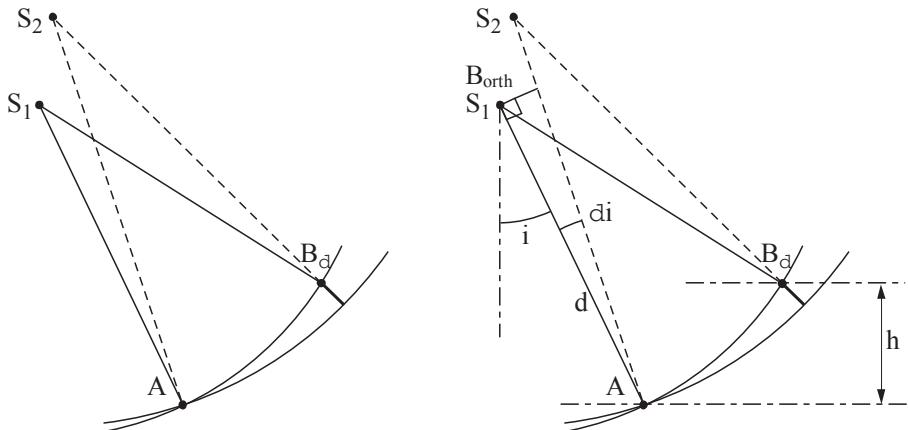


Figure 12.2. Interferometry geometry: transmission and reception antenna successively in S_1 and S_2

In this way, the land is divided into constant elevation slices; each one corresponding to a phase turn of $2\pi r$, similar to a network of “iso-elevation” curves; the elevation corresponding to a phase turn is ambiguity height z_{amb} .

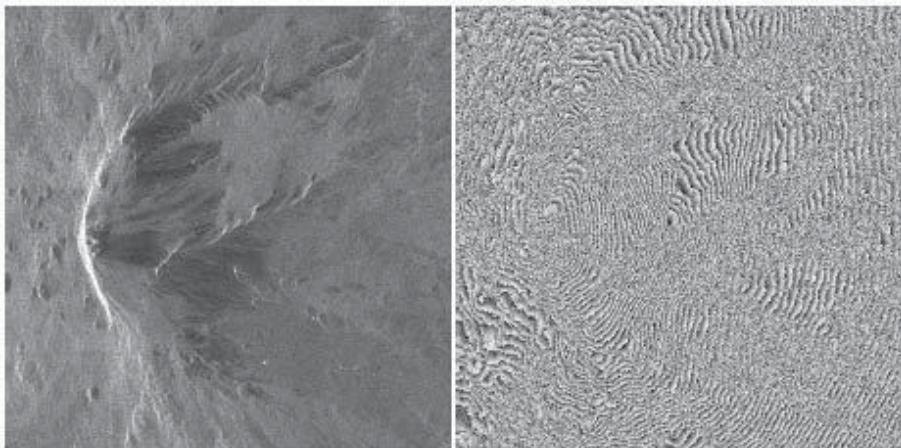


Figure 12.3. Interferogram on Etna obtained by ERS-1:
left, image in elevation; right, interferogram © CNES

12.4.1.1. Limitations

The limitations of this technique are due to the time shifting between acquisitions (this will be analyzed in section 12.5.1), and significant reliefs, creating phase coherence losses (this will be analyzed in section 12.5.2) and to atmospheric artifacts, in the case of non-simultaneous acquisitions (this will be analyzed in section 12.5.3).

The first effect can be eliminated by completing both acquisitions simultaneously or quasi-simultaneously. This is the case of the SRTM mission (see section 3.3.2.4), but also of airborne systems provided with two antennae: one in transmission and reception, the other simply in reception (bi-static system).

The second effect is decreasingly significant, as the range resolution of the system used is good.

12.4.1.2. *Precisions*

Altimetry precisions of the DTM developed by the radar interferometry technique depend on the interferogram's noise level (see section 12.5.4) and generally correspond to an ambiguity height fraction z_{amb} . For ERS we can (on average) "divide" a fringe into 4 or 5. For the SRTM mission, the announced precision corresponds to a fourth of the fringe. Future systems, of better resolution, are expected to improve this value even more. Today, excluding rather costly laser systems, high-resolution airborne radar interferometry is the operational technique that offers the best performances for the completion of DTM and DEM³. There are American and German operational systems, such as IFSARE or Aer-1 (see section 3.6.7), which display decimeter precisions. They are associated with completely automatic treatment chains.

12.4.2. *Land movement measurement: differential interferometry*

We now place ourselves in the case of non-simultaneous acquisitions. If one part of the image's points between the two acquisitions moves (because of landslide, earthquake, volcano swelling, etc.), range to the radar is modified and the phase difference is affected proportionally to radial displacement (toward the radar) $\delta\varphi = \frac{4\pi\delta}{\lambda}$. This measurement is relative. Therefore, it can be perceived only when it affects image points in various ways: global translation of the field would be imperceptible. It is quite straightforward (no stereo effect) and accurate. A half-wavelength movement (28 mm for ERS) toward the radar is translated by a rotation of a phase turn. In theory, following the interferogram noise level and the extension of the effect to be measured, we could observe phase differences as low as $\frac{2\pi}{10}$ or a movement of about 3 mm for ERS. In practice, it is generally admitted that precision is in the order of a centimeter, which still remains spectacular when we know that these measurements are generated by satellite images, taken from distances of hundreds of kilometers. This technique is used by many laboratories for applications such as the study of volcanoes (in particular Etna, Piton de la Fournaise, etc.), earthquakes (Landers, Northridge, Kobe, Izmit, etc.), glaciers, tectonic plates, land settling due to oil or gas extraction, etc.

³ DEM = Digital Elevation Model: it is different from the DTM (Digital Terrain Model) because it describes, in addition, the main underground components, buildings, works of art, forest coverage, etc.

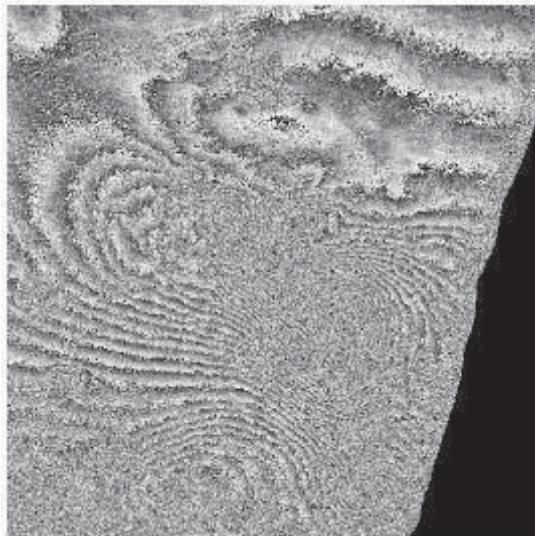


Figure 12.4. Differential interferogram obtained with ERS-1 on the Landers fault in California © CNES

12.4.3. Use of coherence

Empirical coherence degree $d(x, y)$ (equation [12.4]) is a measurement of phase stability between two radar images obtained in neighboring geometric conditions. It is included in interval $[0, 1]$.

A strong coherence translates the fact that two images are identical, including at the level of static speckle distribution. This is only possible if, between the two acquisitions, the land has not changed and geometric conditions are very close.

A weak coherence translates the fact that there has been a change between the two acquisitions, which may have several origins:

- geometric origins: a too significant look angle change causes the coherence to drop. This effect is amplified by the presence of slopes oriented toward the radar (see section 12.5.2). This effect also exists in the case of simultaneous acquisitions;
- changes appearing on land (in the case of non-simultaneous acquisitions);
- water surfaces (lakes, seas, etc.), constantly moving, are always incoherent;
- areas covered with vegetation often have a low coherence. In fact, this depends on the period of time separating the two acquisitions and on the wavelength used. It

is difficult to give an exact description of this dependence because vegetation types vary. However, it is possible to consider that an agricultural plot remains consistent for a week on average for C band (ERS, RADARSAT);

- forests have low coherence because they cumulate rapid growth, movement (for example, under the effect of wind) and volume backscattering whose effects are similar to those of slopes;

- finally, any change on land, even if subtle, affecting the phase, can be detected by coherence measurement: vehicle passage, tilling, plowing, streaming, etc.

Coherence information is completely independent of the image amplitude information and can be used for many applications, such as automatic detection of coastlines or of areas having suffered changes (erosion, flooding), classification into wooded area, non-wooded area, indication of human activity (harvest, plowing, building sites), etc.

The only interferometry product that is sufficiently operational to be commercialized today is *tandem coherence*, provided by the satellite pair ERS-1/ERS-2 (acquisitions taken at 24-hour intervals). Since uniform vegetation remains consistent, it is possible to differentiate it from forests. In a general way, this makes it possible to characterize the land according to its capacity of preserving a stable phase for one day (humid or marshy areas, erosion, etc.).

12.4.3.1. *Implementation*

In practice, the three effects due to reliefs, land movement and loss of coherence are simultaneously present in the interferogram. In particular, it is necessary to be able to eliminate relief effect in order to observe land movements by differential interferometry. More precisely, the recommendations that can be made are the following:

- if the focus is on topography, it is necessary to choose one (or several) image pair(s) with an ambiguity height adapted to the reliefs to be measured (in order to limit losses in geometric coherence: for example, from 50 to 100 m for a moderate relief) and a time period as short as possible (the ideal being simultaneous acquisitions) in order to limit coherence losses due to changes;

- if the focus is on land movements, it is first necessary to choose an interval of time adapted to the phenomenon to be observed (for example, by strictly time-framing an earthquake, or over several years for slow movement). Then, it is necessary to choose an orbital shift as low as possible in order to minimize geometric coherence losses and relief effect. If this is not the case (high number of topographic fringes), an accurate DTM is necessary in order to remove them. Otherwise, we should be satisfied with a rough DTM;

– we retain the same constraints for coherence applications: a time period adapted to phenomena to be observed and an orbital shift as low as possible in order to avoid confusing geometric coherence loss with loss of coherence connected to a state of surface modification.

12.5. Limitations of interferometry imaging

12.5.1. *Temporal change limitations*

In the case of non-simultaneous acquisitions, all the changes happening on land between both acquisitions are liable to modify the statistical distribution of radar signal and therefore its phase. In this case, interferometry technique cannot be applied (except for the detection of these areas by low coherences). The most important causes for coherence loss over time are:

- change of vegetation response, as a result of its growth, after several days, or because of any other event (harvest, plowing, strong wind, etc.);
- perpetual movement of liquid surfaces (seas, lakes, ponds, etc.), with which marshy or unstable areas can be associated;
- land erosion, if the interval of time is significant or if there is a rapid event such as landslide;
- any other change attributable to human activity (for example, we find here phenomena such as the development of parking areas in commercial centers, building sites, deforestation, etc.).

It is necessary to mention the particular case of the forest, which accumulates rapid temporal changes (because of the action of wind and, to a lesser degree, vegetation growth) and geometric coherence losses due to volume backscattering.

12.5.2. *Limitations due to geometry*

It is easy to understand the loss of interferometry coherence due to changes appearing on land between two acquisitions, such as vegetation growth or permanent movement of water. It is more difficult to understand the origin of geometric coherence loss because of look angle variation, which exists even in one pass interferometry. The objective of this section is to describe in detail the origin of a radar image phase and the geometric effects that may alter it.

We have seen that a pixel of a radar image contains, in general, a large number of elementary reflectors (Chapter 5) and that the amplitude and phase of this pixel,

seen from the radar, result from the coherent sum of the response of each elementary reflector, which we have considered as a random reconstruction at the origin of speckle. If two acquisitions are performed in the same geometric conditions and if elementary reflectors have not moved, we obtain the same sum and therefore the same phase. If geometric conditions change, this sum changes so that the phase is then modified. In this way, we have defined the notion of critical baseline (formula [12.7]), for which coherence becomes zero. For ERS, an incidence difference of 0.05° , corresponding to an orbital gap of about 1 km, makes the coherence drop completely (on flat land).

It is inferred that this effect is dependent on the resolution: the larger the resolution cell, the more sensitive the phase difference will be to a look angle variation (an effect analog to antennae directivity because we may consider the resolution cell as an antenna emitting toward the radar). However, the size of the resolution cell strongly depends on look angle θ and on local land slope α . It is all the more significant as the look angle is small and the slope is oriented toward the radar. The loss of coherence is then stronger.

We have the same phenomenon in the case of volume backscattering, for example, in the case of forests. Due to wave penetration, the resolution cell becomes larger (it must then be considered as a small element of volume) and the loss of coherence is more significant for a given look angle variation (regardless of the fact the leaves are moving) (see Figure 12.5).

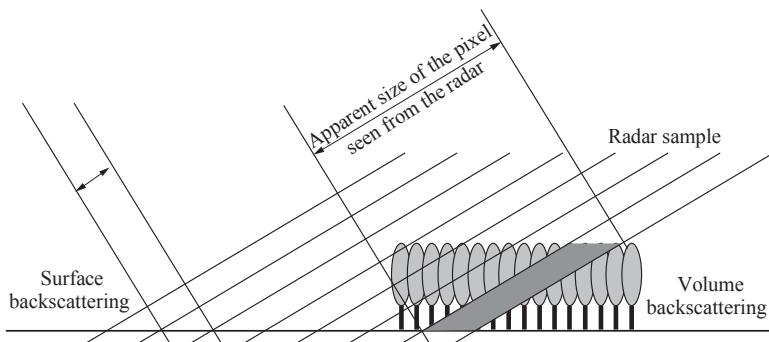


Figure 12.5. The role of vegetation covering in the pixel size on land

There is a different way of explaining this phenomenon, which is based on a frequency approach. When the wave incidence varies slightly, the projection of the carrying frequency on land also varies (see Figure 12.6).

Everything happens as if the central frequency of the emitted wave were varying. A frequency shift between the two images obtained with different incidences appears in this case [GAT 94]. This shift depends on local land slope (α positive for slopes oriented toward the radar) and on $\Delta\theta$, i.e. the look angle variation between both views.

$$\Delta f = \frac{f_c \Delta\theta}{\tan(\theta - \alpha)} = \frac{B_{orth} c}{\lambda D \tan(\theta - \alpha)},$$

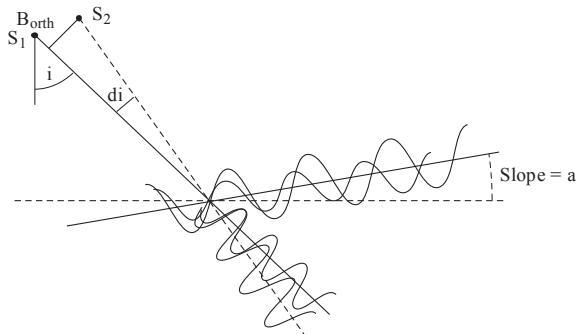


Figure 12.6. Geometric interpretation of coherence loss related to the orientation of the two waves from both sources. Projections on land of carriers initially in phase, are slowly shifted according to land slope

This is clearly in keeping with what is observed when a gap in frequency domain corresponds to a phase incline in temporal domain: these are interferogram fringes. They are all the more tightly knit as orbits are more distant and as land slope is considerable (more significant frequency gap). The common spectrum parts contribute to the formation of interferometry fringes. The rest only brings noise. Some computer software packages can take into account the *a priori* knowledge of land slope in order to filter away these useless parts. Interferometry's limits are deducted from this feature: if the frequency gap is higher than the radar bandwidth (in range), there is no longer a common spectral part and therefore interferometry is no longer possible. We also find that high resolution radar (large bandwidth) is more reliable for slopes: another way of recording this phenomenon is observing the fact that its critical baseline is larger.

Figure 12.8 shows the percentage of common band in the ERS case (at 23°) and for a high-resolution radar (2 m resolution and 45° look angle). Orthogonal baseline B_{orth} is chosen for the case in which ambiguity elevation is 80 m.

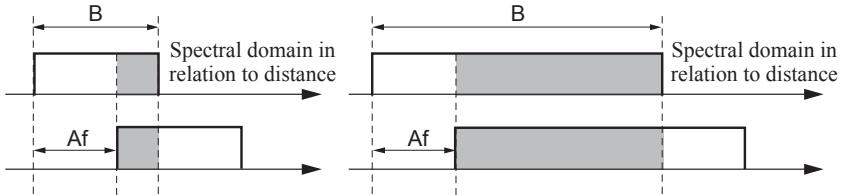


Figure 12.7. Shared spectrum for the two images: for ERS satellites (left) and for a high-resolution satellite (right)

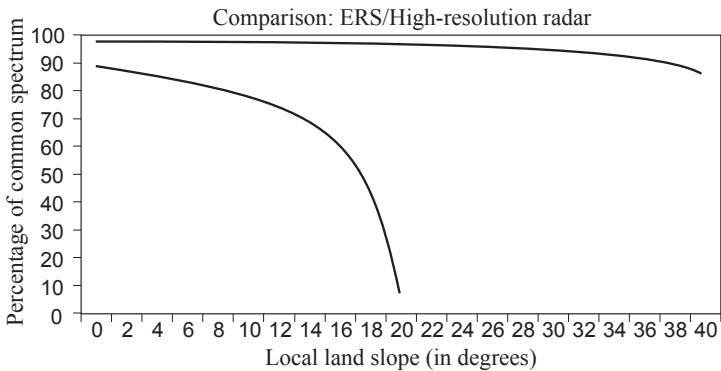


Figure 12.8. Percentage of shared spectrum for both images in relation to land slope: for ERS satellites (curve below) and for a high-resolution satellite (curve above)

12.5.3. Effects of atmospheric propagation

We are placing ourselves in the satellite case, with multi-phase acquisitions, and therefore non-simultaneous. Essentially, interferometry measurement is a range variation measurement (subcentimetric), sometimes because of land displacement, sometimes because of stereo effect. The radar wave crosses twice (round trip propagation) the atmospheric layers (troposphere and ionosphere), which have variable refractive indices. If meteorological conditions are modified, phase differences appear, which disturb those we are trying to measure. Their amplitudes, which are generally low (several centimeters), mainly affect DTM development, where these phase differences are interpreted as belonging to reliefs (via the stereo effect). In this case, the smaller the baseline, the more significant the altimetry errors. In differential interferometry (land movement), these artifacts will hold back final precision (noise). These effects have been pointed out on all interferograms [DUP 97], under the form of low-frequency or high-frequency undulations, of lesser

or higher amplitude. They are attributable to forming clouds, ionosphere bubbles, global humidity variations combined with the presence of reliefs and, more generally, to a variation of the water-vapor content in the troposphere. Consequently, these effects may (unfortunately) take all possible forms (streaks, waveform, clouds, fronts, etc.).

12.5.3.1. Troposphere case

The refractive index in troposphere $n(h)$ is positive and decreases with altitude, as shown, for example, in Bean's atlas (see section 1.1.3). Since the index is higher than 1, propagation velocity is slightly lower than that of light, which is similar to an elongation of wave path.

In addition, the refractive index variation in relation to altitude produces a path curvature. However, it can be shown that this effect, which can manifest itself through minimal variations of the incidence angle, is generally negligible compared to elongation effects. The following integral gives the path elongation for the troposphere crossing according to an average look angle $\bar{\theta}$ and according to co-index $N(h)$:

$$\Delta l_{tropo} = \frac{10^{-6}}{\cos \bar{\theta}} \int_0^{h_{\max}} N(h) dh. \quad [12.8]$$

In relation to an average atmosphere, a differentiation of the Smith-Weintraub formula (see equation [1.11]) giving co-index N takes us to the following expression:

$$\Delta N_{tropo} = -1.45\Delta T + 0.269\Delta P + 4.51\Delta e \quad [12.9]$$

This formula is calculated in standard conditions: $P = 1013 \text{ mb}$, $T = 288.15 \text{ K}$ and a partial water vapor pressure $e = 15 \text{ mb}$. Taking into account this expression in formula [12.8], we can evaluate the path variations in relation to standard conditions.

It can be shown that it is essentially the water vapor content variations (measured in partial water vapor pressure e) that cause these heterogeneities; temperature and horizontal pressure variations are observed only at levels higher than a few dozen kilometers.

Several works [TAR 96] have shown that all the effects, such as fronts, turbulences, convections create index variations (therefore distance and phase

variations), and that the maximum effect is obtained for the cumulus in formation: a few centimeters on average could reach 10 or even 20 centimeters in exceptional configurations. It must be noted that this elongation does not depend on frequency, but it can manifest itself differently in phase, according to wavelength.

12.5.3.2. *Ionosphere case*

Electronic density irregularities in ionosphere layer F (toward 350 km altitude; see Table 1.1) create propagation path shortenings because we have, according to equation [1.10]:

$$\Delta N_{iono} = -\frac{40.28\rho}{f^2}$$

where ρ designates electronic density expressed in electrons per m^3 . Therefore, the ionosphere bias is proportional to the wavelength square. Other works mention distance variations lower than those created by the troposphere for C-band.

Despite the different form of ionosphere irregularities in relation to cumulus-type clouds, it is difficult to differentiate between the two effects. Let us observe, however, the fact that:

- phase difference is proportional to wavelength in ionosphere and inversely proportional in troposphere. Associated distortions are therefore proportional to wavelength square in ionosphere and indifferent to wavelength in troposphere;
- increase in electronic density ρ corresponds to shortening of distances, while water vapor excess creates elongation. However, this is of no help because electronic density ρ decrease, or water vapor impoverishment, create inverse effects.

Finally, it is important to note that ionosphere effects (heterogeneities) and troposphere effects (fluctuations) are of higher amplitude at the end of the day and at dusk than in the morning. A satellite such as RADARSAT, which is phased at 06:00 and 18:00, would then be expected to have a smaller number of atmospheric problems for descending orbits (06:00) than for ascending orbits (18:00).

12.5.4. *Role of signal-to-noise ratio*

It is important to note that there is no speckle noise in an interferogram, because actually the two images that serve to generate it have the same statistical distribution. On the contrary, the signal-to-noise ratio of initial images will have an influence on the interferogram phase; and certainly, loss of temporal or geometric

coherence will entail fringe extinction. Figure 12.9 shows the standard deviation of the interferogram phase clutter and Figure 12.10 shows the corresponding coherence level according to the signal-to-noise ratio of images used and the averaging window size.

It is observed that coherence is systematically underestimated: pure noise gives a coherence of 0.03 with a window of 10 pixels.

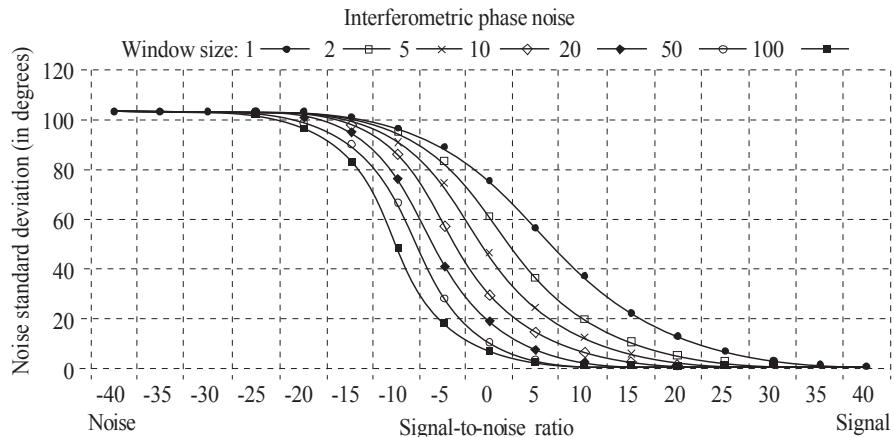


Figure 12.9. Noise phase standard deviation (in degrees) of the interferogram according to signal-to-noise ratio (in dB) of images having served in its elaboration and for averaging windows of 1, 2, 5, 10, 20, 50 and 100 pixels

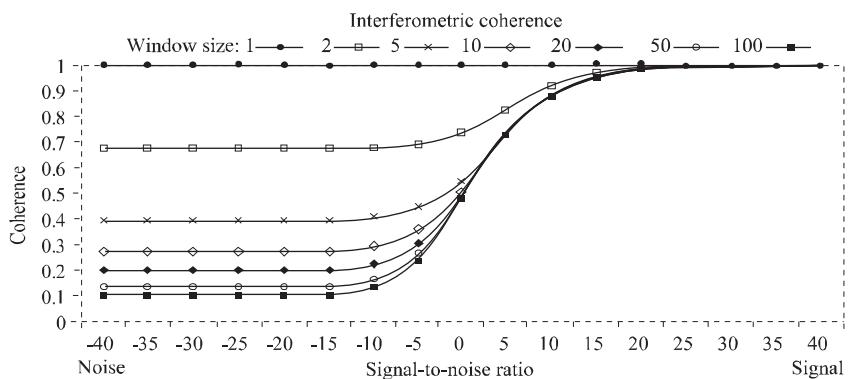


Figure 12.10. Coherence degree of the interferogram according to signal-to-noise ratio (in dB) of images used in its elaboration and for averaging windows of 1, 2, 5, 10, 20, 50 and 100 pixels

Chapter 13

Phase Unwrapping

13.1. Introduction

The various interferometry techniques [REI 86] are faced with the ambiguity of the phase measured modulo 2π , namely the fact that for every pixel P , access is limited to the main value that we will denote hereafter in this chapter by $\varphi(P)$ ¹. To determine relief or a movement field, we will then need to go back to the exact phase value that we will denote by $\Phi(P)$. This stage, which is called phase unwrapping, will consist of finding the right multiple k of 2π such that:

$$\Phi(P) = \varphi(P) + 2k(P)\pi \quad \text{where } \varphi(P) \in [-\pi, \pi[$$

In one dimension, the wrapped phase looks like a seesaw signal. In two dimensions, the measure modulo 2π turns a continuous model into a network of fringes, the edges of which (brutal transitions from 0 to 2π) only depend on the origin of phases on the complex circle. This phenomenon is illustrated in Figure 13.1.

Chapter written by Emmanuel TROUVÉ and Henri MAÎTRE.

¹ For notational simplicity, $\varphi(P)$ will in this chapter stand for the interferogram phase that in the previous chapter had been denoted by $\delta\varphi(P)$. As for $\delta\varphi(P)$ itself, it will from now on represent the phase variation between two pixels in the image.

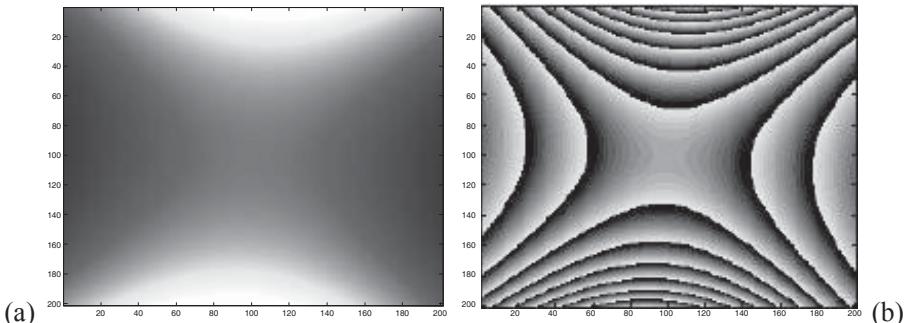


Figure 13.1. Left: unwrapped phase. Right: the same phase wrapped

13.1.1. Corrected phase shift

The two-dimensional phase signal, which in the interferogram is known as modulo 2π , is spatially sampled according to pixel sizes obtained in azimuth and range after initial processing. If the Nyquist criterion is observed in both directions, sampling will have to be performed at twice the maximal signal frequencies for each direction. The precise phase difference between two adjacent samples therefore has to be lower than half period π . If P_1 and P_2 are two neighboring pixels as in the 4-connectivity², we have:

$$|\Delta\Phi(P_1, P_2)| = |\Phi(P_2) - \Phi(P_1)| < \pi \quad [13.1]$$

This assumption makes it possible to define corrected phase shifts $\varphi_c(P_1, P_2)$ based on the difference between main phases $\delta\varphi(P_1, P_2) = \varphi(P_2) - \varphi(P_1)$, according to the rule:

$$\varphi_c(P_1, P_2) = \begin{cases} \delta\varphi(P_1, P_2) & \text{if } |\delta\varphi(P_1, P_2)| < \pi \\ \delta\varphi(P_1, P_2) + 2\pi & \text{if } \delta\varphi(P_1, P_2) \leq -\pi \\ \delta\varphi(P_1, P_2) - 2\pi & \text{if } \delta\varphi(P_1, P_2) \geq \pi \end{cases} \quad [13.2]$$

² On a square grid, 4-connectivity is the dual of 8-connectivity. In 4-connectivity, a point's only neighbors are two adjacent pixels in its row and two other adjacent pixels in its column. In 8-connectivity, a point has four additional neighbors, consisting of adjacent pixels in the diagonals. On discrete grids, connectivity is important for ensuring the topological properties of detected shapes (see, for example, [SCH 94]).

The corrected phase differences along both axes make up a vectorial field denoted by (φ_x, φ_y) , where:

$$\varphi_x(i, j) = \varphi_c((i, j), (i+1, j)) \quad \text{and} \quad \varphi_y(i, j) = \varphi_c((i, j), (i, j+1)).$$

Then, $\varphi_c(P_1, P_2)$ will denote either of the corrected discrete derivatives φ_x or φ_y , depending on the direction of the vector P_1P_2 .

If assumption [13.1] is valid all across the image, unwrapping is easily obtained by propagation from a reference pixel P_0 . Then, the exact phase difference between pixel P_1 and P_0 is given by summing up the corrected shifts along any path $C_{P_0 \rightarrow P_1}$:

$$\Phi(P_1) - \Phi(P_0) = \sum_{P \in C_{P_0 \rightarrow P_1}} \varphi_c(P, P'), \quad [13.3]$$

where P' denotes the successor of P when P describes C . The field of the corrected phase shifts (φ_x, φ_y) has conservative circulation, as the sum along an arbitrary closed path is zero:

$$\sum_{P \in C_{P_0 \rightarrow P_0}} \varphi_c(P, P') = \Phi(P_0) - \Phi(P_0) = 0. \quad [13.4]$$

In SAR interferometry, condition [13.1] is virtually never satisfied over the entire image. The sequential unwrapping in equation [13.3] therefore leads to different results, depending on the path we take, and to discontinuities arising in the unwrapped image. These inconsistencies of 2D phase unwrapping have been presented and discussed by, among others, Ghiglia [GHI 87] and Goldstein [GOL 88].

13.1.2. Residue detection

Equation [13.4] can help us detect the singularities that are responsible for these typical 2D phase unwrapping discontinuities. We consider a basic closed edge made up of four adjacent pixels. The sum of the corrected phase differences along such boundaries is necessarily 0 , 2π or -2π [BON 91]. The points for which the sum is non-zero are called residues that may be negative or positive, according to the sign of the result for one direction of the fixed path. We can detect residues over the entire image and create what we will hereafter call a residue image, as illustrated in Figure 13.2.

These residues play a decisive part in phase unwrapping: their presence reveals inconsistencies of the main value phase, which trigger completely wrong results where local methods are used, or systematically biased ones if we follow global approaches such as least squares methods. Automatic unwrapping methods require that the number and influence of residues be reduced beforehand. The two standard approaches used to this end consist of:

- filtering the interferogram to reduce phase noise. This can lead to dramatic results (see Figure 13.5) in areas affected by reversible disturbances involving a partial decrease of coherence;
- masking the areas where geometric phase signal cannot be reconstructed, i.e., in areas struck by irreversible disturbances such as coherence loss (water, strong temporal evolution, etc.) or aliasing³ in the presence, for example, of too significant reliefs or discontinuities caused by fold-over.

These two often complementary approaches require some preprocessing to make sure that interferograms obtained in harsh conditions (large baseline, strong relief, etc.) are *unwrappable*.

³ Aliasing is a defect occurring where a signal is sampled at a frequency below the lower limit imposed by Shannon's theorem.

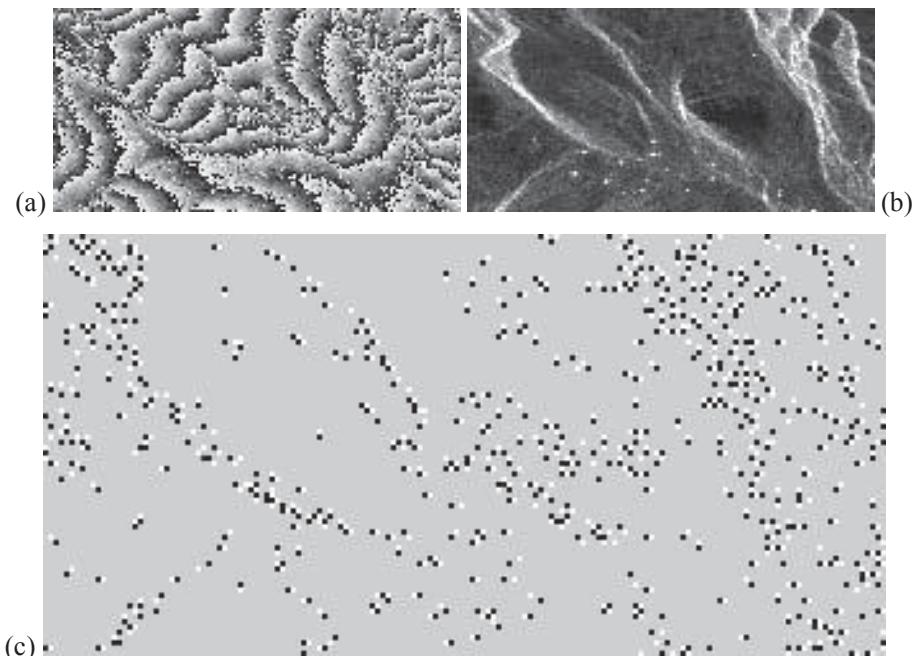


Figure 13.2. Residue examples: (a) wrapped phase and (b) amplitude of the same mountainous area of 80×160 pixels © CNES; (c) residues detected over (a), in black and white, according to their sign (zoom twice). We see that residues accumulate in the folded areas that can be identified by their high amplitude in image (b)

13.2. Preprocessing of InSAR data

This section is devoted to a set of preprocessing techniques with the goal of extracting from interferometric data the robust information that phase unwrapping algorithms necessitate.

We are faced with two problems:

- a detection problem where we try to characterize the presence or absence of the geometric phase signal; and
- an estimation problem in which we try to measure the phase gradient based on which exact phase will be reconstructed.

A solution to both of these problems consists of estimating local 2D frequencies and assigning a confidence measure to this estimator. The approach, which is described in section 13.2.1, provides sufficient information to use global

unwrapping methods on interferograms with few discontinuities. By contrast, local methods cannot be used in uneven areas unless the estimation side is deepened by interferogram filtering to de-noise the wrapped phase (section 13.2.2). In addition, the detection side has to be supplemented, especially by exploiting amplitude data, to identify irreversible disturbances (section 13.2.3).

13.2.1. Analysis of local frequencies

If the phase derivative corresponds to the *instantaneous frequency* of a signal, phase gradient estimation can be regarded as a spectral analysis problem over small-sized windows [SPA 95]. Phase signal modeling by a complex 2D sinusoid makes it possible to rely on standard algorithms such as the periodogram [KAY 81] or MUSIC (*multiple signal classification*) [STO 89]. However, the implementation of these methods is fraught with difficulties [CAR 94] related to calculation time, scant number of samples, etc. Therefore, we will have to find direct estimation algorithms, particularly based on autocorrelation matrix properties of the complex phase signal. Below we will review principles of the algorithm enabling the capture at every point of a local frequency estimate and a confidence measure associated with it [TRO 96b]. We will then propose a multi-scale implementation that improves estimation performance independently of fringe width.

13.2.1.1. Modeling

If in a multi-look interferogram, we denote a noise-free interferometric phase by β and estimated phase by φ , the phase signal we seek is affected by an additive noise η such that:

$$\varphi = \beta + \eta \pmod{2\pi}$$

The probability density of noise $p(\eta)$ is provided by distribution $p(\varphi|\beta)$ according to the number of looks and coherence D (equation [12.5]). To preserve phase continuity, the interferogram is taken in the complex form of unit module $s(m,n) = e^{j\varphi(m,n)}$. Observed signal s_b is then marred by a multiplicative noise $b(m,n)$:

$$s_b(m,n) = s(m,n) \cdot b(m,n) = e^{j\beta(m,n)} \cdot e^{j\eta(m,n)}$$

Assuming that the broadly stationary noise is spatially decorrelated and decorrelated from the signal, the noise autocorrelation function (ACF) is:

$$\gamma_b(m, n) = E\left(e^{j\eta(m,n)}e^{-j\eta(0,0)}\right) = \begin{cases} 1 & \text{if } (m, n) = (0, 0) \\ K = \left|E\left(e^{j\eta(0,0)}\right)\right|^2 & \text{otherwise.} \end{cases}$$

Parameter K is a function of the noise probability density and varies between 1 in the absence of noise and 0 in the case of noise with uniform distribution on $[-\pi, \pi]$. If the phase signal is locally approximated using the deterministic model of a complex 2D sinusoid of frequencies (f_x, f_y) : $s(m, n) = e^{j2\pi(mf_x + nf_y)}$, the ACF of the noisy signal $\gamma_{s_b}(i, j, k, l)$ is written (see [TRO 96b]) according to the ACF of the γ_s model and to differences $m = i - k$, $n = j - l$, as follows:

$$\gamma_{s_b}(m, n) = E\left(s_b(i, j) \cdot s_b^*(k, l)\right) = K_{\gamma_s}(m, n) + (1 - K)\delta_{m, n} \quad [13.5]$$

we deduce power spectral density \mathcal{D} of the noisy signal:

$$\mathcal{D}\{s_b\}(u, v) = K \cdot \delta(u - f_x, v - f_y) + 1 - K$$

where (u, v) describe frequency space $[-0.5, 0.5] \times [-0.5, 0.5]$. The peak we seek $\delta(u - f_x, v - f_y)$ is reduced by factor K and immersed in white noise of power $1 - K$. The presence of a significant noise makes it tricky to estimate pair (f_x, f_y) by seeking the spectral density maximum. The standard periodogram method that would work optimally in terms of maximum likelihood for a sinusoid affected by additive white Gaussian noise is therefore hard to apply to SAR interferograms [MAR 98].

13.2.1.2. Direct estimation method

By simplification, we consider square windows of size $D_s \times D_s$ to make up the signal vector by row concatenation [PRA 75] and of size $D_e \times D_e$ ($D_e > D_s$) to estimate

its autocorrelation matrix (ACM), assuming ergodicity. According to equation [13.5], ACMs Γ_s and Γ_{Sb} of pure and noisy signals verify relation:

$$\underline{\Gamma}_{Sb} = K \cdot \underline{\Gamma}_s + (1 - K) \cdot I \quad [13.6]$$

where I is the identity matrix of size $D = D_s^2$. By introducing frequency vector $\underline{e}(f_x, f_y)$ such that:

$$\underline{e}(f_x, f_y)^\dagger \begin{bmatrix} 1 \\ e^{j2\pi \cdot f_x} \\ e^{j2\pi \cdot 2f_x} \\ \dots \\ e^{j2\pi(D_s-1)f_x} \\ \hline \dots \\ e^{j2\pi \cdot f_y} \\ e^{j2\pi(f_x+f_y)} \\ \dots \\ e^{j2\pi((D_s-1)f_x+f_y)} \\ \hline \dots \end{bmatrix}$$

we can write the ACM of pure signal Γ_s as:

$$\underline{\Gamma}_s = \underline{e}(f_x, f_y) \cdot \underline{e}(f_x, f_y)^\dagger$$

We thus find the following property related to the Toeplitz structure of ACM blocks:

- each row L_j within a block of rows is deduced by multiplying the preceding row by $a_x = e^{j2\pi \cdot f_x}$:

$$\underline{L}_{j+1} = a_x \cdot \underline{L}_j \text{ where } j = j_o + kD_s, j_o \in [1, D_s - 1], k \in [0, D_s - 1] \quad [13.7]$$

– the rows within each block can be deduced by multiplying the preceding block by $a_y = e^{j2\pi f_y}$:

$$\underline{L}_{j+D_s} = a_y \cdot \underline{L}_j \text{ where } j \in [1, D_s(D_s - 1)] \quad [13.8]$$

According to equation [13.6], relations [13.7] and [13.8] hold for Γ_{Sb} , except where the equalities involve diagonal elements. For each of the two unknowns a_x and a_y , we thus have $D_s(D_s - 1)$ relations between rows that lack 2 elements, i.e. $N_e = D_s(D_s - 1)(D_s^2 - 2)$ equations. We can then estimate factor a_x (or a_y , respectively) minimizing the square error $\|\underline{v}_2 - a_x \cdot \underline{v}_1\|^2$ (or, respectively, $\|\underline{w}_2 - a_y \cdot \underline{w}_1\|^2$) where the vectors \underline{v}_1 and \underline{v}_2 (respectively \underline{w}_1 and \underline{w}_2) are formed starting from N_e elements in the Γ_{Sb} rows that satisfy relation [13.7] (or [13.8], respectively):

$$a_x = e^{j2\pi f_x} = \frac{\underline{v}_1^\dagger \cdot \underline{v}_2}{\|\underline{v}_1\|^2} \quad \text{and} \quad a_y = e^{j2\pi f_y} = \frac{\underline{w}_1^\dagger \cdot \underline{w}_2}{\|\underline{w}_1\|^2}$$

Frequency \hat{f}_x (respectively, \hat{f}_y) is given by the argument of scalar product $\underline{v}_1^\dagger \cdot \underline{v}_2$ (respectively $\underline{w}_1^\dagger \cdot \underline{w}_2$). The two frequency components can thus be estimated via a constant number of operations, which is proportional to D_s^4 . The gain obtained over methods such as MUSIC [TRO 96a] makes the estimation of local frequencies compatible with the size of InSAR data.

13.2.1.3. Confidence measure

As irregular fringes may occur in some areas of an interferogram, the hypotheses on which the estimation of local frequencies relies are not always verified. To avoid using wrong frequencies, the validity of every estimate has to be checked. To this

end, the above approach supplies two measures of confidence that differ from each other in nature:

- one of them, c_a , is related to the least squares estimation algorithm: residual square errors $\|\underline{v}_2 - a_x \underline{v}_1\|^2$ and $\|\underline{w}_2 - a_y \underline{w}_1\|^2$ are normalized and combined according to pair (\hat{f}_x, \hat{f}_y) . This confidence measure depends on the noise level that affects Γ_{sb} coefficients. It shows whether the algorithm can deliver a reliable estimate;

- the other measure, c_m , has to do with model validity. The frequency vector $\underline{e}(f_x, f_y)$ corresponding to the sinusoid must be the eigenvector related to the highest eigenvalue of Γ_{sb} (estimation principle by MUSIC algorithm). By checking whether solution $\underline{e}(\hat{f}_x, \hat{f}_y)$ complies with this property, we can *a posteriori* verify model validity.

As both these measures are sensitive to various disturbances (noise, model breach, etc.), they are merged via a pessimistic operator such as the harmonic mean to form a confidence measure in the estimation: $c_e = 2c_m c_a / (c_m + c_a)$. This measure is then used either directly as a tool to detect areas that cannot be estimated at original interferogram resolution, or as part of a multi-scale approach to merge several estimates.

13.2.1.4. Multi-scale implementation

The performance of the above algorithm depends on estimation window length as compared with fringe size. For a typical size-9 signal vector ($D_s = 3$) and an ACM estimated over 49 realizations ($D_e = 9$), the estimation will be robust to noise on relatively narrow (up to 12 pixel wide) fringes. However, where high noise occurs, the wider fringes would require either greater estimation windows or a decrease in resolution accompanied by low-pass filtering. To observe the local nature of stationarity and minimize calculation time, we will favor the second solution and apply the estimation algorithm and confidence measure to several scales.

This approach poses two difficulties:

- aliasing of the narrowest fringes as we construct scales Λ_l , where a pixel corresponds to $l \times l$ pixels in the original scale Λ_1 . If coherence measure d is available, we will prefer to downscale by truncating the 2D Fourier transform by signal $de^{j\varphi}$ (see Figure 13.3). A pseudo-coherence d^l will thereby be preserved at

every scale and reveal the regions of uncertain phase as well as high-frequency regions suppressed by truncation;

– a merger of redundant or even conflicting information originating from different scales. A merger strategy consists of combining at each scale l confidence measure c_e^l and coherence d^l to obtain a global confidence c_g^l . Every estimate will then correspond to hypothesis $h^l = (\hat{f}_x^l, \hat{f}_y^l)$ that is checked against other hypotheses via a compatibility coefficient $r(h^l, h^{l'}) = c_g^l \cdot c_g^{l'} \cdot (1 - \delta(h^l, h^{l'}))$ where $\delta(h^l, h^{l'})$ is the distance between 2D frequencies. Optimal hypothesis h^{opt} is the one that maximizes the sum of compatibilities over the entire set of hypotheses. The final frequency and associated confidence are then calculated by averaging values obtained at different scales, weighted by compatibility coefficients $r(h^{opt}, h^l)$. We thus refine the most likely estimate, while eliminating the influence of conflicting estimates.

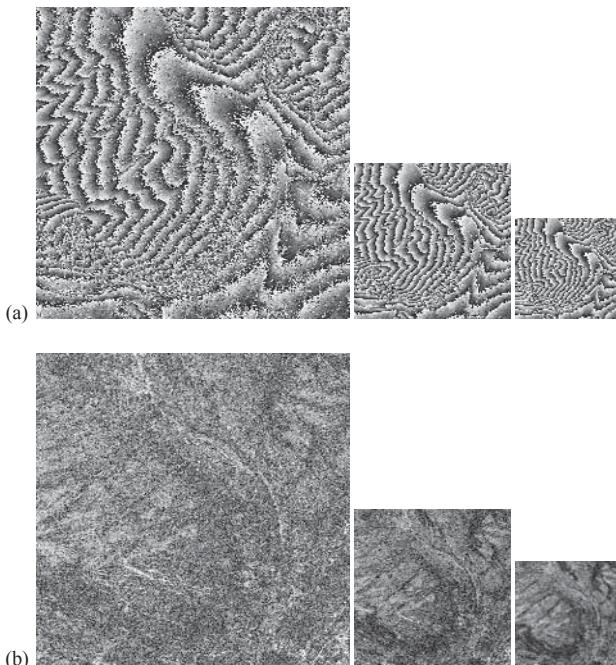


Figure 13.3. Multi-scale data construction: (a) phase, (b) coherence. Original resolution (© CNES), plus twofold and threefold reductions

By adding typically 3 or 4 downsampled versions to the original interferogram, multi-scale frequency analysis makes it possible to analyze fringes different in width with the same effectiveness in terms of noise. Local frequencies are then employed either in an immediate phase in which the wrapped phase will be filtered (section 13.2.2), or directly as a phase gradient as part of an unwrapping method. Likewise, the confidence measure can be used in the stage in which non-unwrappable areas are detected (section 13.2.3), or directly to weight unwrapping by the least squares method (section 13.3.2).

13.2.2. Interferogram filtering

The various processing methods aimed at improving interferogram quality include:

- preprocessing techniques such as spectral shift [GAT 94] that rely on complex initial data and are an integral part of the interferogram generation method (see Chapter 12);
- post-processing techniques relying on the three pieces of data that users are generally provided, i.e., images of phase φ and empirical coherence d , resulting from complex M -look processing (see equation [12.4]) and the amplitude image A produced by averaging the intensity over the $2 \times M$ pixels available in the initial resolution⁴.

We review the two most standard SAR data filtering approaches and then present a post-processing technique leading to a considerable reduction of phase variance and empirical coherence.

13.2.2.1. Two different approaches

Whether radiometric or interferometric data are concerned, the filtering techniques proposed in radar imagery belong more often than not to either of the following approaches:

- a signal/image approach based on filtering out useless frequencies in the Fourier domain. The interferogram generator software (FRINGE software) provided by ESRIN appears to favor this particular approach, as most de-noising operations are performed by band-pass filtering either directly on the spectrum, or in the spatial domain using convolutions with the corresponding cardinal sine [PRA 94];

⁴ We cannot consider here that we are dealing with a $2 \times M$ -look image, as speckle realizations between the two images are *a priori* strongly correlated (basic hypothesis of interferometry).

– a statistical approach based on speckle modeling [GOO 75] and helping to reduce estimate variance by averaging over a certain number of realizations. The ideal case is the genuine M -look intensity treatment. However, unless looks or independent images are available in sufficient numbers, these techniques average neighborhood supplied samples, assuming an ergodic stationary signal. Coherent spatial averaging over M samples of product $z_1(i, j) \cdot z_2^*(i, j)$, where z_1 and z_2 are the two originally complex images, is the standard solution known as complex multi-looking [ROD 92]. It is applied in CNES software [MAS 93a] (DIAPASON software) that has provided the data used for illustrations in this chapter. Some speckle filters based on this approach have been extended to interferometric data. Among them are the Lee filter [LEE 94a], which nonetheless requires that the phase be previously unwrapped, and the Lopes vectorial filter [LOP 93a], based on minimization of the mean square error (MMSE). Since only neighborhood-collected samples are available, this filter yields a weighted spatial averaging that does not seem to match up to the results obtained in intensity estimation by MMSE filtering of multi-date SLC data [BRU 94].

Tested on real data [TRO 96a], the two approaches yielded visually and in residue numbers very close results, except for the spectral aliasing case where filtering is accompanied by sub-sampling. In practice, both approaches have the same problem of phase non-stationarity in narrow fringes. These natural variations contradict the assumption of a neighborhood of samples sharing the same phase and disturbed by noise alone. Considering that the statistical approach leads to slightly better results and is more flexible to implement, we find it will be preferable to adapt a complex averaging filter to local signal characteristics.

13.2.2.2. Fringe pattern compensation

Knowing the fringe pattern provided by the 2D local frequency makes it possible to overcome the above difficulty. Let (\hat{f}_x, \hat{f}_y) be the local frequency as estimated at point P . With P as origin, phase evolution in its neighborhood is illustrated by the following plane model:

$$\varphi_{\text{mod}}(i, j) = 2\pi \left(i\hat{f}_x + j\hat{f}_y \right) \pmod{2\pi}.$$

These variations can be compensated by forming the difference $\varphi(i, j) - \varphi_{\text{mod}}(i, j) \pmod{2\pi}$ at every point in the image. Adaptability to the fringe pattern is thus achieved: we have a neighborhood in which the phase no longer varies except as a result of noise and local variations with respect to the model.

Complex averaging over a growing number of assumedly independent samples thus becomes possible. If the registered initial complex data $z_1 = A_1 e^{j\phi_1}$ and $z_2 = A_2 e^{j\phi_2}$ are available, phase and coherence at pixel P can be estimated as the argument and module of a complex correlation over a window \mathcal{F} :

$$\hat{d}(P) e^{j\hat{\varphi}(P)} = \frac{\sum_{(i,j) \in \mathcal{F}} A_1(i,j) e^{j\phi_1(i,j)} A_2(i,j) e^{-j\phi_2(i,j)} e^{-j\varphi_{\text{mod}}(i,j)}}{\sqrt{\sum_{(i,j) \in \mathcal{F}} A_1^2(i,j)} \sqrt{\sum_{(i,j) \in \mathcal{F}} A_2^2(i,j)}}.$$

If all we have is phase, coherence and amplitude images which are the result of a first averaging of M -looks, we can reconstruct the sum of M initial samples by approximating the two amplitudes $A_1(i,j)$ and $A_2(i,j)$ via average $A(m, n)$:

$$\sum_{(i,j) \in \mathcal{F}} z_1(i,j) \cdot z_2^*(i,j) \approx M A^2(m,n) \cdot d(m,n) \cdot e^{j\varphi(m,n)}$$

where (i, j) denotes the pixels at initial resolution, while (m, n) denotes those of the three available data. Filtering then goes back to the statistical approach and extends the complex averaging to a window W with fringes being compensated for:

$$\hat{d}(P) e^{j\hat{\varphi}(P)} = \frac{\sum_{(m,n) \in W} A^2(m,n) \cdot d(m,n) \cdot e^{j\varphi(m,n)} \cdot e^{-j\varphi_{\text{mod}}(m,n)}}{\sum_{(m,n) \in W} A^2(m,n)}. \quad [13.9]$$

We thus obtain a filtered phase $\hat{\varphi}$, with a strongly reduced variance, and an extension of coherence measure d , which is more regular than the initial coherence (see Figure 13.4).

The filtering window size may be fixed or may vary to provide better adaptability to noise level. The number of single-look samples that will be necessary to bring back phase variance φ below an acceptable threshold is a function of level of coherence D between the two initial images [LI 90]. We can determine the order of magnitude of the number of samples to be averaged using the Cramer-Rao bound [ROD 92]:

$$\sqrt{E\left(\left(\hat{\phi} - E(\hat{\phi})\right)^2\right)} = \frac{1}{\sqrt{2M}} \frac{\sqrt{1-D^2}}{D}$$

by using coherence D estimate provided by coherence image $d(P)$. We can for instance use windows of varying sizes $W(d(P))$, from the four nearest neighbors (4NN) for a well correlated area up to 9×9 windows for the noisiest areas. Results thus obtained are illustrated in Figure 13.5 and summarized in Table 13.1 in the form of a reduced number of residues over $1,024 \times 1,024$ pixel images, from which Figures 13.4 and 13.5 are extracted. The rate of irreducible residues, which is in the order of 2% to 3% of total number of pixels, is due to fringeless areas, discontinuities, etc. that are inaccessible to spectral estimation and that we will try to detect in the next stage.

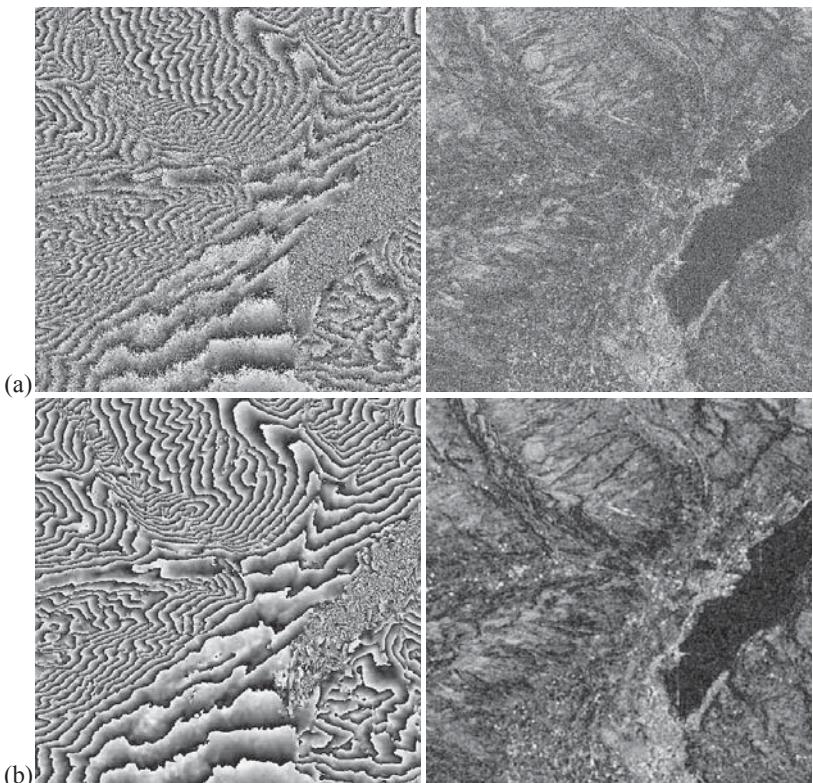


Figure 13.4. Fringe and coherence measure filtering: (a) initial phase and coherence, 10-looks, 512×512 -pixel images, © CNES; (b) filtering according to equation [13.9]; 5 x 5 window for phase (left) and 3 x 3 window for coherence (right)

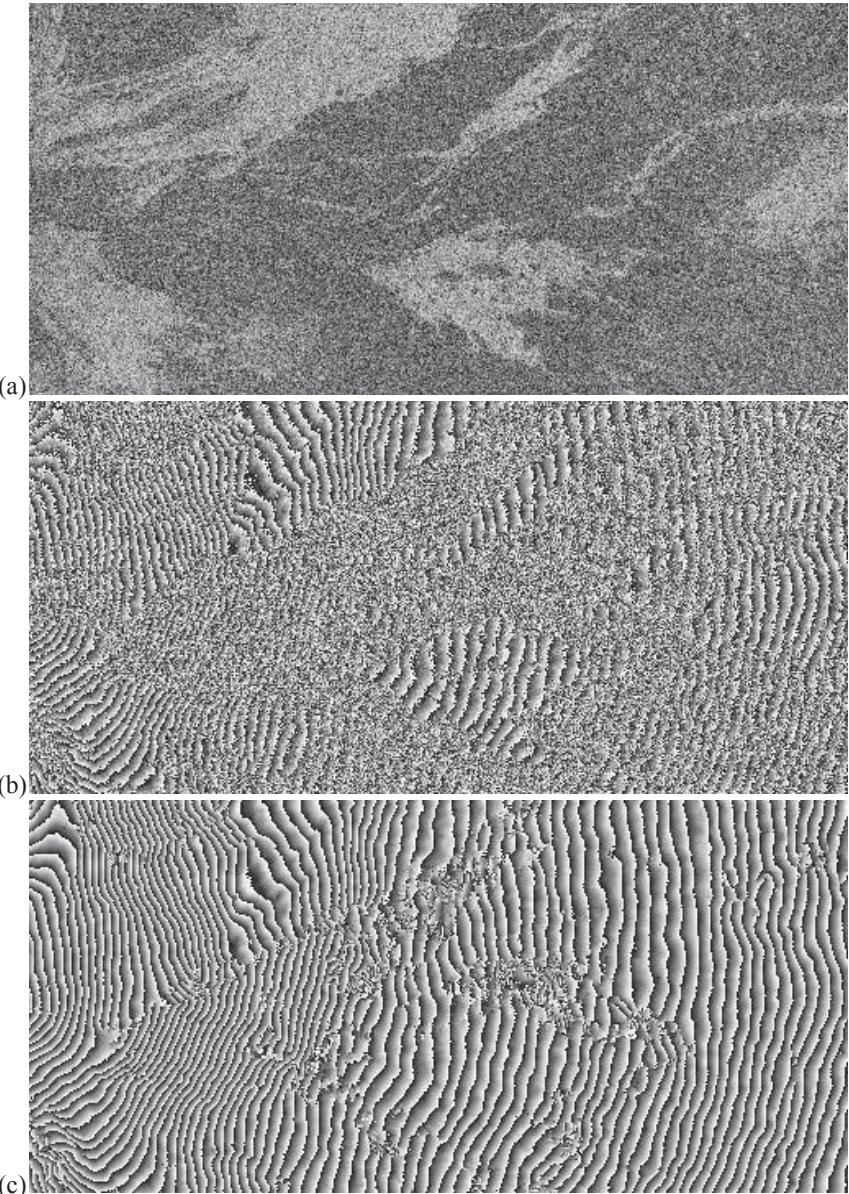


Figure 13.5. Fringe filtering: (a) coherence; (b) initial phase; 8-looks, 256 x 512-pixel image © CNES; (c) interferogram filtered through windows with size adapted to noise level from 4 NN to 9 x 9 depending on coherence. Remaining disturbances are due to zones out of the spectral estimation reach: absence of fringes or relief discontinuities

Scene	Number of looks	Initial image	After filtering					
			4NN	3x3	5x5	7x7	9x9	$W(d(P))$
Berne	2 x 10	15.8	4.9	3.3	1.9	1.6	1.5	1.7
Etna	2 x 4	20.2	7.2	4.6	2.9	2.6	2.5	2.7

Table 13.1. Residue point percentage depending on filter size. Increasing window size favors the disappearance of residues. The remaining 2-3% is due to fringeless zones; no significant improvements can be obtained below this limit. Notation 4NN shows that the filtering window is made up of just the central pixel and its 4 nearest neighbors. Window $W(d(P))$ varies in size depending on the empirical level of coherence at the considered point

13.2.3. Detection of non-unwrappable areas

This third intermediate stage is aimed at identifying certain structures that are responsible for irreversible disturbances. To this end, we can adopt a data fusion approach within a contextual classification made via a Markov model (see Chapter 7). The classes describe the various disturbances that are for the most part related to terrain elements where the surface or slope is incompatible with interferometry.

Depending on the content of the processed scene, we will introduce:

- classes describing decorrelated areas: water planes (lake or river), fields and forests that have undergone temporal evolution, etc.;
- classes describing discontinuity areas: slopes with scant sampling, fold-over areas; and
- a complementary class that identifies unwrappable phase areas.

The data are selected so that they have discriminatory value with respect to the disturbances that we are seeking to detect. What we initially have is the coherence image (Figure 13.6a), which reveals the weak coherence areas, and amplitude image (Figure 13.6b), which indicates the relief slopes facing the radar. For a first noise reduction in these two images, we use the filter defined in equation [13.9] for coherence and a speckle filter (see Chapter 5) for amplitude. Specific data are sometimes necessary to obtain a finer characterization of some classes. We thus can introduce a confidence measure linked to an estimation of local frequencies (Figure 13.6d) to identify the areas that have no geometric phase and those where fringes can still be estimated despite considerable noise. We will also resort to a directional contrast measure (Figure 13.6e), as derived from road detectors (see Chapter 8), to detect the finest over-brightness arising from compression areas.

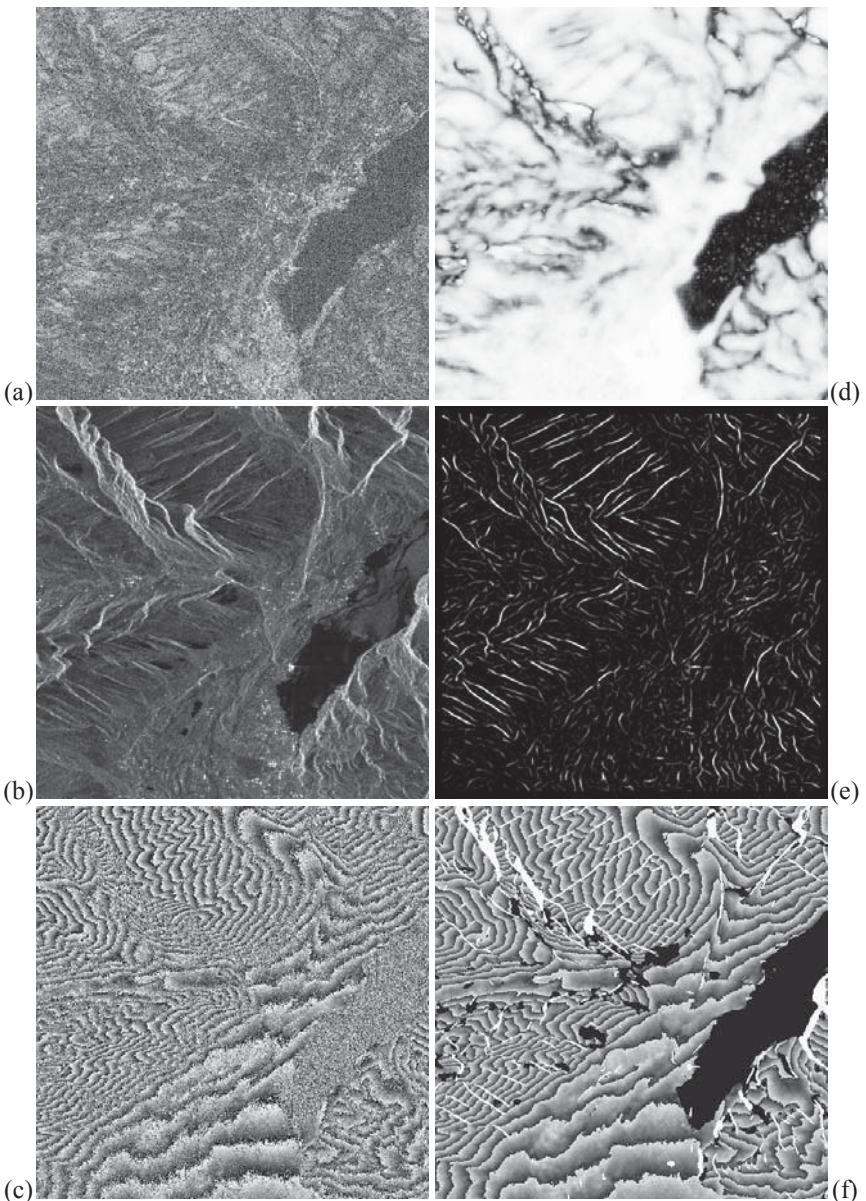


Figure 13.6. (a) Coherence, (b) amplitude, (c) phase: initial data © CNES, (d) confidence in frequency estimation, (e) directional contrast measure, (f) filtered interferogram and non-unwrappable areas (black: decorrelated; white: discontinuous), as detected based on the measures (a), (b), (d) and (e) above (see color plate section)

Each class has its characteristic conditional distributions observed in images which show the corresponding disturbances. Fusion is achieved in the energy that is attached to the data and consists of the sum of elementary potentials translating at each site the agreement between image information and the different classes [DES 95]. The *a priori* probability is given by the standard Potts potential expressing agreement among the classes assigned to sites that belong to the same set. To preserve the thin structures during regularization, this model can be supplemented by adding an external field that takes into account directional information [TUP 96]. Classification is obtained by minimizing energy through simulated annealing.

This type of approach makes it possible to obtain the expected results, i.e., covering most residues persisting after filtering by classes that reveal the disturbances and keeping most of the interferogram in the unwrappable class (Figure 13.6f).

13.3. Phase unwrapping methods

The advent of SAR interferograms has revived interest in two-dimensional phase unwrapping. The numerous methods that have been advanced fall into four broad categories:

- local methods where phase is unwrapped by propagation from one point to the next. The problem raised by residue connection and several different solutions are presented in section 13.3.1;
- regional methods that try to bring decisions down to the level of elementary regions. Without going into detail, we can cite several papers that addressed the issue in terms of fringe segmentation [GIE 87, LIN 92, PER 95], while others looked at the pairing of regions following various methods of partitioning [TOW 89, JUD 92, TAR 96];
- global methods that try to find a solution on the image as a whole. Among them are techniques based on the first Green identity [FOR 96] that are presented in section 13.3.3 below and the least-squares methods described in section 13.3.2;
- finally, there are a number of halfway methods relying on local unwrapping rules while seeking global solutions via approaches such as cellular automaton [GHI 87, SPI 91] or Markov fields [DUP 95]. These methods, which involve a heavy calculation cost, are generally implemented on parallel computers.

The two most widespread types of methods are described in this section [TRO 98, ZEB 98]. For a comprehensive review of phase unwrapping in radar interferometry, see Ghiglia and Pritt [GHI 98].

13.3.1. Local methods

Local unwrapping methods are based on a simple rule analog to Cauchy's theorem for analytic functions [GOL 88]:

*The sum of the corrected phase shifts φ_c
along a closed path is zero if and only if the numbers
of positive and negative residues contained within it are equal.*

[13.10]

In other words, any closed path constructed during unwrapping must enclose an equal number of positive and negative residues. To this end, we try to connect residues of opposite signs with cuts that unwrapping paths are not supposed to cross. We thus make sure that the sum of the corrected phase shifts is zero along any authorized path. This is no guarantee that the unwrapping is accurate, but at least it will be coherent. Accuracy depends on how pairing of the cuts is selected and where the cuts are placed with respect to disturbances generating the residues (see Figure 13.7).

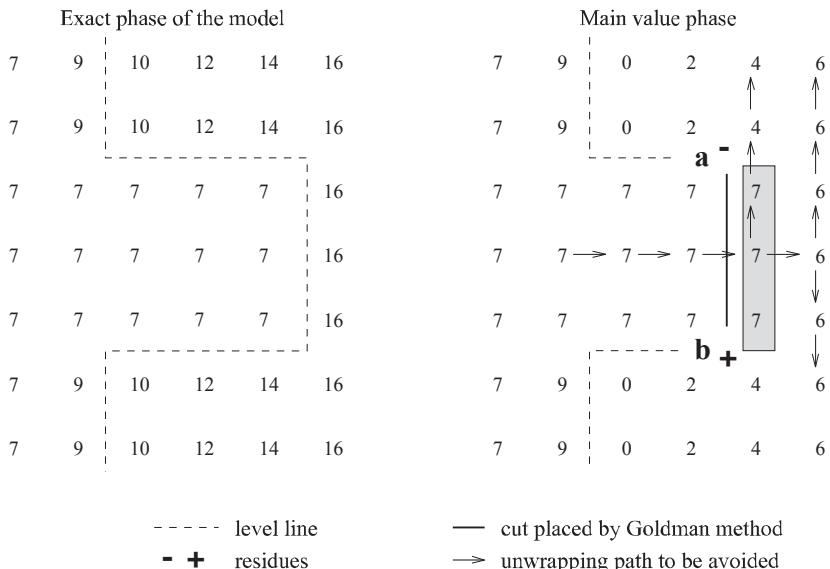


Figure 13.7. Cut positioning between residues *a* and *b* to prevent unwrapping beyond the ghost line. Here the phase is codified modulo 10. Pixels in the grayed area will be unwrapped to 17 rather than 7. Still, the error is confined within the area comprised between the cut and the real level line 10

The main problem with these local methods then lies in grouping residues into components that are neutralized by either balancing charges, or reaching the edge of the image. Two main types of approach may be used to do so: positioning the cuts in the shape of lines that connect the residues and finding a mask where connected components have to be balanced and correspond to the disturbances. In both cases, we can considerably simplify the problem by first trimming down the residue number with a filter that matches fringe pattern and noise level (see section 13.2.2).

13.3.1.1. *Placing the cuts*

The cuts must first connect residues belonging to the same disturbance. For residues caused by an impulse type noise, pairing in terms of distance is certainly the safest and most direct method [GOL 88, HUN 89].

Where continuity breaks are responsible for residues, phase folding wipes out the level lines that correspond to changing 2π multiples (see Figure 13.7). The endpoints of this ghost line are detected as residues [PRA 90b]. Placing a cut to connect them makes it possible to avoid propagation of incorrect unwrapping. The mistake will be confined to the area between the cut and the exact position of the line. However, two residues may lie far apart and therefore be difficult to connect to each other: the odds are that we form incorrect cuts by connecting them to closer residues and end up with an unwrapping that will no longer match reality. Ways to improve positioning of the cuts were proposed notably by relying on the amplitude image to favor grouping of residues along structures with the same radiometry [PRA 90a].

13.3.1.2. *Finding a mask*

Where we have a high density of noise-related residues (for example, in decorrelated areas), positioning cuts is not only costly but also useless, as it generally leads to a lot of cuts where the inside cannot be unwrapped. A faster approach consists of creating a mask to directly cover the residues belonging to non-unwrappable areas. Depending on interferogram characteristics, we can find the mask by relying exclusively either on the phase image, by thresholding, for example, the maximum of the three second order discrete derivatives [BON 91], or on phase and empirical coherence, by thresholding a filtered coherence following local unwrapping [DER 95], or on the three interferometric pieces of data following some significant classification work (see section 13.2.3).

Forming a mask based on amplitude and coherence data has the advantage of regrouping the residues according to physical criteria that match the disturbances. Most importantly, we will thus accurately bring together residues which are far apart, but belonging to the same compression area at the origin of a major discontinuity. This type of mask, however, needs to be refined so that connected components that do not touch the edge cover positive and negative residues in equal

numbers. This operation shown in Figure 13.8 can be conducted by inserting residues of a filtered interferogram into an original mask and analyzing charges of the different connected components of the image thus formed. Isolated residues and mask components cluster progressively according to a distance criterion and until charge balance is reached over all components of the image as a whole. The result can then be unwrapped by propagation along any path within the unmasked areas and then by interpolation into the forbidden areas as well.

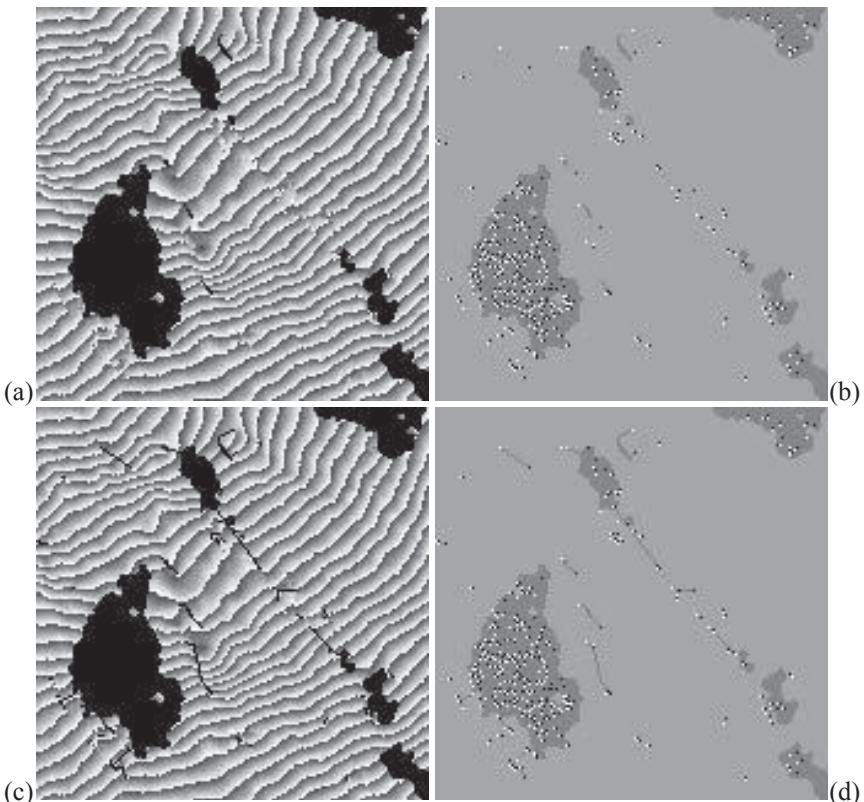


Figure 13.8. Canceling the charge of a mask's connected components; (a) – (c): the mask before and after pairing; (b) – (d): positions of the residues (black/white, according to the sign) and of the mask (in dark gray) before and after pairing

13.3.2. Least squares

A very different approach to the problem of two-dimensional phase unwrapping consists of seeking a global solution via the least squares method. The only

unambiguous piece of information that is *a priori* available is the phase gradient $\nabla\varphi = (\varphi_x(i, j), \varphi_y(i, j))$ that is calculated on the interferogram by the corrected phase shift rule (see equation [13.2]). For an image of size $N \times N$ (square for simplicity), we try to determine the unwrapped phase $\Phi(i, j)$, the discrete gradient of which minimizes the square error:

$$\begin{aligned} E &= \|\nabla\Phi - \nabla\varphi\|^2 \\ &= \sum_{i=1}^{N-1} \sum_{j=1}^N (\Phi(i+1, j) - \Phi(i, j) - \varphi_x(i, j))^2 \\ &\quad + \sum_{i=1}^N \sum_{j=1}^{N-1} (\Phi(i, j+1) - \Phi(i, j) - \varphi_y(i, j))^2 \end{aligned} \quad [13.11]$$

The solution to equation [13.11] corresponds to that of the discrete Poisson equation with a Neumann condition (i.e., that the gradient component normal to the image edge is zero) [GHI 89]:

$$\Phi(i+1, j) + \Phi(i-1, j) + \Phi(i, j+1) + \Phi(i, j-1) - 4\Phi(i, j) = \varrho(i, j) \quad [13.12]$$

where:

$$\varrho(i, j) = \varphi_x(i, j) - \varphi_x(i-1, j) + \varphi_y(i, j) - \varphi_y(i, j-1) \quad [13.13]$$

This equation holds for pixels located inside the image, i.e., $i \in (2, N-1)$ and $j \in (2, N-1)$. On the sides and at the angles, the central pixel factor becomes 3 and 2, respectively, (compared with 4 in equation [13.12]). We recognize the left hand term as being discrete Laplacian $\nabla^2\Phi$ of the phase we seek. As for the source term ϱ , it is the discrete version of divergence of vector field $\nabla\varphi = (\varphi_x, \varphi_y)$.

13.3.2.1. Considering disturbances

Our previous interpretation of equation [13.12] makes it possible to reveal the inaccuracy of the least squares solution where residues are present. By analogy with electromagnetism [MOR 53], the vectorial field $\nabla\varphi$ is decomposed into a sum of two fields: an irrotational field $S_{ir} = \nabla u$ that is equal to the gradient of a scalar potential $u(x, y)$ and a zero divergence field S_d . The field S_{ir} is a conservative

circulation field, according to Stokes' theorem and corresponds to the regular part of the phase gradient. The presence of residues, however, leads to the existence of a component S_d that disappears as the source term is calculated (equation [13.13]). Then the solution of the Poisson equation is given by $\Phi(x, y) = u(x, y)$, where the residual square error is $E = \|S_d\|^2$. We see that around residues, the difference between the gradient of the solution $\nabla\Phi$ and the phase difference field $\nabla\varphi$ drops to $1/r$, where r is the distance to the residue point [TAK 88a].

Unwrapping by the least squares method then tolerates the presence of isolated residues that only cause limited errors within their neighborhood. By contrast, with higher residue densities, the solution based on corrected phase shifts is systematically biased (underrated) compared with the exact model. This global error arises from the bias that affects gradient estimation from the wrapped phase, according to equation [13.2] [LOF 97].

When residues are the result of reversible disturbances such as noise in partially decorrelated areas, we can avoid this type of error by adopting either of the following strategies:

- filtering the interferogram so as to considerably reduce the influence of noise in the unwrappable areas (see section 13.2.2) while keeping the corrected phase shifts as gradient estimator;

- using another, unbiased and more robust against noise, gradient estimator. We can for instance rely on the local frequency estimator described in section 13.2.1: $\nabla\varphi = (2\pi\hat{f}_x, 2\pi\hat{f}_y)$. These estimates are unbiased, as they are obtained from autocorrelation coefficients (see equation [13.5]) $\gamma_{s_b}(m, n) = K \cdot e^{j2\pi(mf_x + nf_y)}$ in the form:

$$\hat{f}_x = \frac{1}{2\pi} \arg \left[\sum_{m,n} \gamma_{s_b}(m+1, n) \cdot \gamma_{s_b}^*(m, n) \right]$$

The direct use of local frequencies to calculate the source term (equation [13.13]) also makes it possible to considerably reduce noise-related phase fluctuations without any filtering (see Figure 13.9). Finally, robust estimation methods have specifically been developed [LYU97] to estimate directional phase derivatives.

When the residues are the result of irreversible disturbances such as discontinuities or missing geometric phase signal, weighting has to be introduced to

suppress influence by these zones as the final solution is calculated. We then minimize the error:

$$E_p = \sum_{i=1}^{N-1} \sum_{j=1}^N w(i,j) (\Phi(i+1,j) - \Phi(i,j) - \varphi_s(i,j))^2 + \sum_{i=1}^N \sum_{j=1}^{N-1} w(i,j) (\Phi(i,j+1) - \Phi(i,j) - \varphi_y(i,j))^2 \quad [13.14]$$

The source term of the partial derivatives equation [13.13] then becomes:

$$\begin{aligned} \varrho_p(i,j) &= w(i,j)^2 (\varphi_x(i,j) + \varphi_y(i,j)) \\ &- w(i-1,j)^2 (\varphi_x(i-1,j) - w(i,j-1)^2 \varphi_y(i,j-1)) \end{aligned} \quad [13.15]$$

Weights $w(i,j)$ may originate from different mask research techniques. Figure 13.9 illustrates a least squares unwrapping on the famous differential interferogram of the 1992 earthquake in Landers, California [MAS 93b]. The gradient used in equation [13.15] is obtained by local frequency estimation and the binary weighting by thresholding of the associated confidence measure [TRO 98].

13.3.2.2. Minimization algorithms

The effectiveness of the least squares method also depends on the existence of digital analysis techniques that make it possible to calculate solutions on large images ($N \geq 1,024$) within a reasonable time interval.

The unweighted Poisson equation solution is commonly obtained by switching to the spectral domain [SPA 93]. Using the discrete Fourier transform (DFT) implies data periodization that cannot be overlooked [PRI 94]. A rigorous solution will consist of constructing a $2N \times 2N$ image by mirror reflection along the two directions [TAK 88b]. A similar version consists of using the discrete cosine transform (DCT) without having to duplicate the data [GHI 94].

In order to solve the weighted least squares solution, we have resorted to iterative techniques that increase calculation cost. The slow convergence of method types such as Jacobi, Gauss-Seidel or iterative over-relaxation (IOR) makes them unusable for large images. In the circumstances, we have to use either multi-grid methods [PRI 95a] [PRI 95b], or conjugate gradient methods that take few iterations to converge provided

that the system is well conditioned [GOL 89]. Convergence can be accelerated by a preconditioning that may involve, for example, resolving the unweighted system at every iteration [GHI 94]. To yield the result seen in Figure 13.9, this method typically requires some 20 iterations on 1,024 x 1,024-pixel images to converge. The iteration time mainly depends on the time needed to calculate the unweighted solution, i.e., a few seconds by switching to the spectral domain.

13.3.3. Green's functions

Another way of finding the minimum of relation [13.11] square error involves the use of the first Green's identity [FOR 96]:

$$\iint_S dS (\Phi \nabla^2 g + \nabla \Phi \cdot \nabla g) = \oint_C dc \Phi \frac{\partial g}{\partial n_S},$$

where g is the corresponding Green's function for the investigated problem, Φ is the unwrapped phase we try to find, S is the integration surface, C is its contour and n_S is an outward directed unit vector orthogonal to the contour's tangent.

For an infinite 2D domain (which is the case widely adopted in other works), function g is equal to:

$$g(r - r') = -1/2 \cdot \ln|r - r'|,$$

which leads to the following expression of the unwrapped phase according to the wrapped phase:

$$\Phi(r') = \iint_S dS \nabla \varphi(r) \cdot \nabla g(r - r') - \oint_C dc \Phi(r) \frac{\partial g(r - r')}{\partial n_S}. \quad [13.16]$$

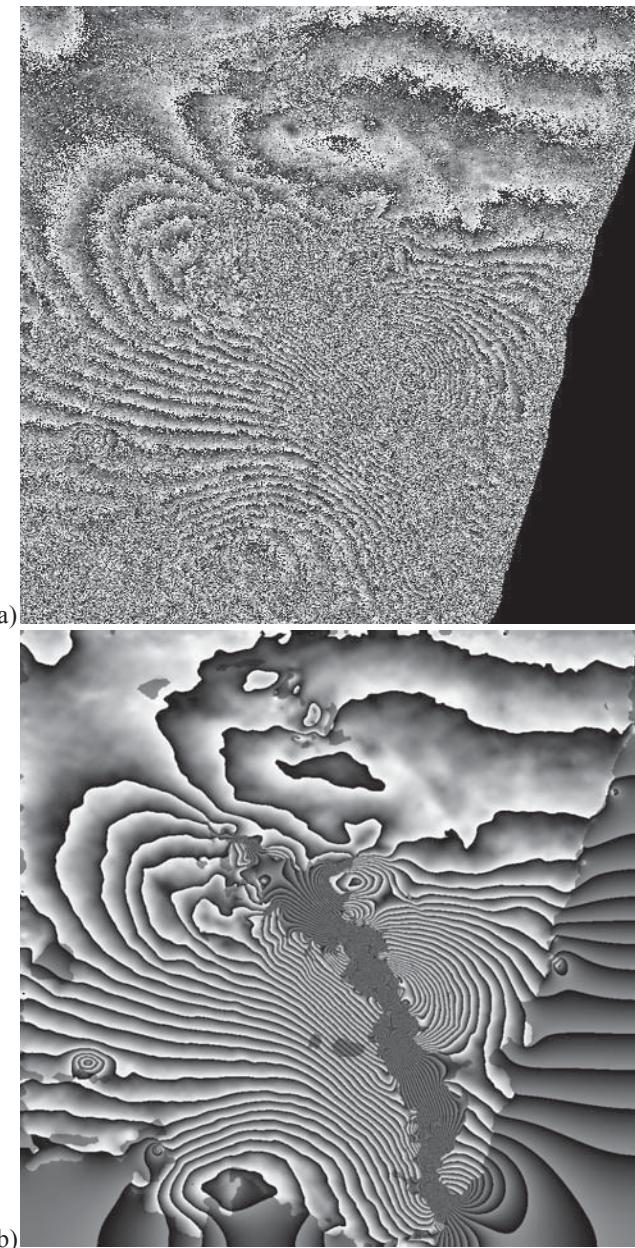


Figure 13.9. (a) 2,048 x 2,048-pixel differential interferogram on California's Landers fault © CNES; (b) phase unwrapped by the weighted least squares (wrapped for comparison, mask in dark hue) (see color plate section)

The presence of term Φ on both sides of the equation leads to resolving it iteratively starting with a zero value or estimate of Φ on edge C and then recalculating it with the Fredholm equation of the second kind:

$$\frac{1}{2}\Phi(r_c) = \iint_S dS \nabla \varphi(r) \cdot \nabla g(r - r_c) - \oint_C dc \Phi(r) \frac{\partial g(r - r_c)}{\partial n_S},$$

as the line integral does not contain point r_c itself. Such technique has been shown to converge in a dozen iterations [FOR 97].

Under another solution proposed in [LYU 98], the iterative resolution can be avoided by using a Green's function expression in the form of an unwrapping of the Helmholtz eigenfunctions. For a two-dimensional square domain with side a , g functions are written:

$$g(r, r') = \frac{1}{a^2} \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \frac{F_{mk}(r) F_{mk}(r')}{v_{mk}^2},$$

where $F_{mk}(r)$ and v_{mk}^2 are the eigenfunctions and eigenvalues of the Helmholtz equation:

$$\nabla_2 F_{mk}(r) + v_{mk}^2 F_{mk}(r) = 0.$$

We infer that:

$$F_{mk}(x, y) = \cos(\xi_m x) \cos(\xi_k y) \text{ and } v_{mk}^2 = \xi_m^2 + \xi_k^2, m, k \in [1, \infty],$$

where $\xi_k = k\pi/a$. The contour integral of equation [13.16] is zero for these functions, as the integrated term is also zero on the contour, and equation [13.16] comes down to:

$$\Phi(r') = \iint_S dS \nabla \varphi(r) \cdot \nabla g(r, r')$$

which can be integrated without iterations. Such a technique leads to significant savings on calculation time.

In this formalism, solutions were also proposed to explicitly take into account paired residues during fringe unwrapping, which leads to accurate solutions if the residues are adequately paired. The mirror image method is adopted in this situation [LYU 99].

This fringe unwrapping method was also extended to circular domains [LYU 00].

13.3.4. Conclusion

Overall, unwrapping methods depend on two essential data:

- a robust phase gradient estimation: as the corrected phase shifts are very quickly disturbed by noise, it is generally necessary to filter the wrapped phase in order to cut down the amount of residues, or use more complex measures such as local frequencies; and
- a detection of the areas where the above data are unavailable: to identify the disturbances that wipe out the geometric phase signal, we have to take into account all information that can be collected or measured on the three initial images.

Once these two elements have been obtained by adequate processing, it is preferable that we choose:

- a local method if we are in a position to completely solve the problem of residue grouping. This is a simple way of finding an accurate solution that respects model discontinuities;
- a global method if the opposite is true. While this approach may often be more complex to implement, it helps obtain a coherent, nearly exact solution, without requiring that we remove all ambiguities: a partial detection of the areas where the phase gradient is corrupt will be enough to minimize their influence on final results.

This page intentionally left blank

Chapter 14

Radar Oceanography

14.1. Introduction to radar oceanography

This chapter is devoted to the mechanisms of electromagnetic wave backscattering on the sea surface and the SAR imaging of this surface. The focus is on the backscattering mechanisms due to sea surface wavelets, the modulation of which makes it possible to observe many phenomena and therefore deserves special attention. Given their operating frequency, radars indeed, whether with a real or synthetic aperture, can only detect surface details (roughness) that are very fine compared with the overall image size. This roughness, however, while associated with surface winds and the intrinsic motion of the surface, is modulated by phenomena on a far vaster scale (swell, internal waves, bathymetry, etc.). All such phenomena that can be imaged by radars are shown in Figure 14.1. The interaction of all these elements and the fact that the phenomenon observed by the sensor is in uneven motion lead to a complex image signal (i.e., a non-Gaussian and often non-stationary and non-linear signal), the analysis of which thus becomes all the more exciting.

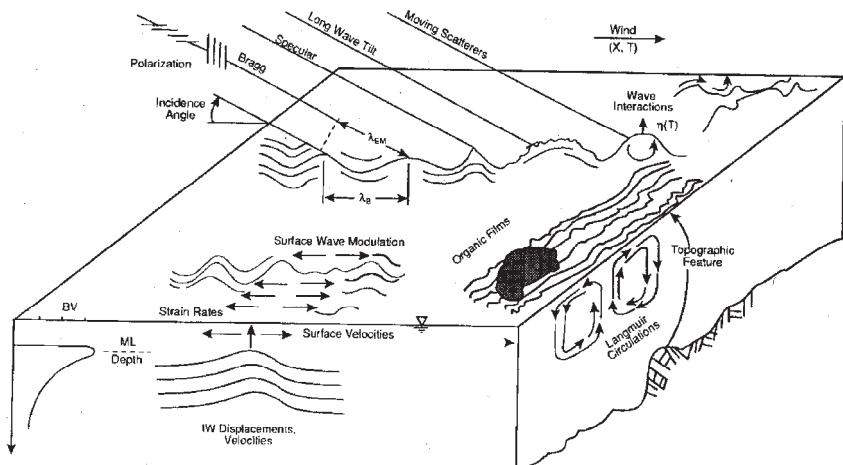


Figure 14.1. Illustration of surface roughness modulation by mesoscale marine phenomena (according to Backscatter, 7, 1, 1996)

These phenomena are clearly represented in Figure 14.2 [ROD 98]. The surface phenomena (swell, internal waves) to be characterized seem to be largely mesoscale. The methods to detect and characterize these phenomena are beyond the purpose of this work, but can be found in full detail in [GAR 07]. For further information, see the references mentioned in the following descriptions:

- high resolution spectral analysis methods (autoregressive modeling and time-frequency representation) have made it possible to study at length the variations that occur in swell direction and wavelength [CAR 91, GRA 97]. The resolution cell is in the order of a square kilometer in this case;
- time-scale (wavelet) analysis has been used for the semi-automatic detection of internal waves on radar images and possibly a physical characterization of the phenomenon [ROD 97, ROD 99].

As we have just seen, the ocean surface roughness (wavelets) to which the SAR is sensitive is modulated by large scale structures. Three mechanisms – tilt, hydrodynamic modulation and velocity bunching – have been identified as being responsible for such modulation. In addition to these wave imaging mechanisms, some effects tend to degrade the images. The azimuth component of the displacement velocity of the target (scatterer) introduces an error into the phase history estimation with respect to the reference signal. This leads to an image blurring called azimuth smearing. The radial component of acceleration leads to the very same effect of decreasing image contrast and filtering small details. Since these

problems have to do with the synthetic aperture process, adequate refocusing should be able to improve image contrast and quality. What may be the best focusing to adopt is still a moot question. However, since (non-linear) velocity bunching and azimuth smearing have been identified on the SAR spectrum, these phenomena can be introduced without going into considerations about the sea that are not readily accessible to readers unfamiliar with the topic. The surface of the sea and theories of electromagnetic wave reflection on the sea surface are next described in this chapter. Finally, a SAR specific application is approached, including notably a brief account of the SAR spectrum inversion method [HAS 91] and an illustration of how the imaging mechanism is modeled [ING 99a] to measure seabed bathymetry by radar imaging.

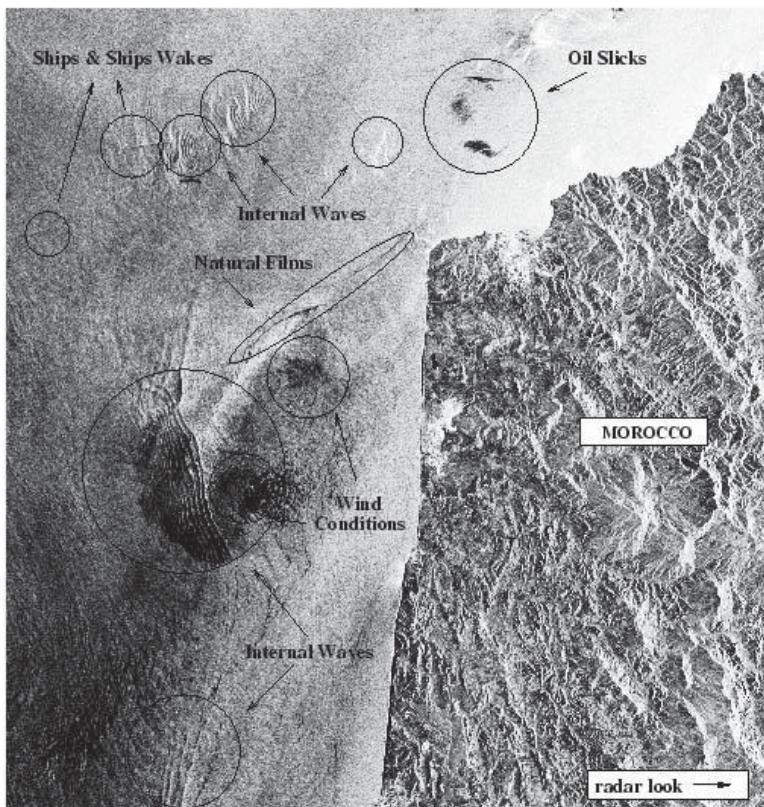


Figure 14.2. ERS-1 image acquired at 11:05, September 10, 1993 (frame 2889, orbit 11261) near the Strait of Gibraltar. The image shows a typical set of surface modulations (© ESA). The results are extracted from [ROD 98]

14.2. Sea surface description

14.2.1. Division of the sea spectrum into different scales

Fluid mechanics equations on a liquid element and in the absence of external forces lead to the wave propagation equation (Navier-Stokes equation) [MCC 73]:

$$\rho \left(\frac{\partial V}{\partial t} + V \cdot \nabla V \right) = -\nabla p + \tau \nabla^2 V \quad [14.1]$$

where V is the velocity, p is the pressure, ρ is the density of water and τ is the water surface tension. The term $V \cdot \nabla V$ is a non-linear term that will be neglected in the linear propagation theory presented below. In open waters and deep waters (without involving surface to bottom height), the surface is comprised of all of the solutions of equation [14.1] and is described as an infinite and continuous sum of independent oscillators. The height of the sea surface $z(x, t)$ is given by:

$$z(x, t) = \int_{-\infty}^{+\infty} \zeta(k) e^{j\varphi(k)} e^{j(kx - \omega t)} dk = \int_{-\infty}^{+\infty} e^{j(kx - \omega t)} dZ(k),$$

where x is a 2D vector and $\zeta(k)$ is the Fourier transform of $z(x, t)$. The possibility of finding solutions to equation [14.1] has led to a division of sea spectrum waves into three classes, depending on the term that is neglected in the equation: capillary waves, the wavelength of which is lower than 5 mm and which are characterized by a dispersion relation $\omega^2 = \frac{\tau}{\rho} k^3$, gravity waves having wavelengths higher than 10 cm and a dispersion relation $\omega^2 = gk$, where g is the weight acceleration, and a transition class termed gravity-capillary waves, the dispersion relation of which can be written by combining the two previous dispersion relations into one: $\omega^2 = gk + \frac{\tau}{\rho} k^3$ [NEU 66]. As pointed out above, it is this latter class of waves that are responsible for the backscattering phenomena at the sea surface. Sea surface slopes are spatial derivatives of $z(x, t)$. Similarly, the velocities and acceleration are temporal derivatives and therefore can be expressed using the linearly filtered surface Fourier coefficients:

$$w(x, y, t) = \int_{-\infty}^{+\infty} D(k) \zeta(k) e^{j\varphi(k)} e^{j(kx - \omega t)} dk$$

The transfer function $D(k)$ takes on values depending on the context.

Velocity along the vertical axis Oz , for example, is deduced from:

$$\frac{dz(x, y, t)}{dt} = \int_{-\infty}^{+\infty} \zeta(k) e^{j\varphi(k)} (-j\omega) e^{j(kx - \omega t)} dk$$

For a monochromatic wave, the propagation velocity along the other axes (displacement velocity of a point located at the same height) is the phase velocity of this wave:

$$\frac{dx}{dt} = v_\varphi = \frac{\omega k_x}{|k|}$$

If monochromatic waves are thus superimposed, the propagation speed of the surface is the sum of the propagation speeds of the various waves (linear propagation theory).

The projection of this speed along an axis of a direction vector \vec{u} such that $\vec{u} = (0, \sin \theta, -\cos \theta)$ (e.g., the radar incident radius) can be linearly deduced from the Fourier coefficients of the surface. The transfer function is then:

$$D(k) = \sin \theta \frac{\omega k_y}{k} - j\omega \cos \theta \quad [14.2]$$

In everything we have described so far, we implicitly assumed that every wave (oscillator) on the sea surface was independent in phase and amplitude from any of the other waves (linear theory). There is, however, a description that implies a non-linear coupling of the sea surface waves taking into account $V \cdot \nabla V$ in equation [14.1]. By taking this term into consideration, we can account for why the third order moment of the slopes is non-zero, as Longuet-Higgins found [LON 63, LON 82]. Non-linear effects explaining this departure as well as wave breaking have been studied by several authors without achieving a convincing model [ZAK68, KRA 94, HAS 61]. However the SAR signature of these effects is important solely for costal areas. Assuming a surface description by independent oscillators implies a Gaussian statistic of that surface, which is a useful hypothesis when we calculate the SAR spectrum and bispectrum. Sea surface waves can be described quantitatively using the significant wave height that is defined as:

$$H_s = 4\sqrt{E(z^2(x,t))} = 4\sqrt{\int_{-\infty}^{+\infty} E(\zeta(k)^2) dk}$$

This quantity is key when describing a sea state (calm or rough) and provides an idea of just how important SAR transformation linearities are.

14.2.2. Reflection of electromagnetic waves from the sea surface

Wright [WRI 66, WRI 68] has proposed writing the backscattering section of the sea in the unique form:

$$\sigma^o = \frac{4 \cdot k^4}{\pi} \left| g(\theta) \int_{-l_x/2}^{l_x/2} \int_{-l_y/2}^{l_y/2} \gamma e^{2jky \cos \theta} dx dy \right|^2 \quad [14.3]$$

where

$$g_{vv}(\theta) = R_v \sin^2 \theta + \frac{1}{2} T_v^2 \left(1 - \frac{\epsilon_0}{\epsilon} \cos^2 \theta \right) \quad \text{and} \quad g_{hh}(\theta) = R_h \sin \theta,$$

where T_v and R_v are the Fresnel coefficients of the refraction-reflection from the planar surface of the same dielectric. The surface portion illuminated by the wave is within the range $-l_x/2, l_x/2$ to $-l_y/2, l_y/2$. The temporal and spatial average over this surface portion makes it possible to rewrite equation [14.3], by introducing the spectrum of sea surface elevation $D(x,y)$, as:

$$\sigma^o = \frac{4 \cdot k^4}{\pi} |g(\theta)|^2 S_D(0.2k \cos \theta)$$

In order to calculate σ^o , it is often considered that $S_D(-0.2k \cos \theta)$ follows the Phillips equilibrium law [PHI 77], or to put it another way, that:

$$S_D(k) = \frac{1.17}{2} 10^{-2} k^{-4} \quad [14.4]$$

The result is proportional to the surface spectrum of the wavelength ($0.2k \cos \theta$). If this coefficient is zero or very weak, there may not be any backscattering. Due to this backscattering proportionality, the phenomenon is called a Bragg backscattering resonant phenomenon and $2k \cos \theta$ is called the Bragg wavelength.

14.3. Image of a sea surface by a real aperture radar

The image we obtain by mapping the backscattering section is thus the image of a real aperture radar (RAR). RAR imaging is generally considered to be a linear process. What we are dealing with is a modulation by the sea surface of σ^o around a mean. We will therefore talk of modulation transfer functions (MTFs), which of course are linear transfer functions between a modulation and the sea. The RAR MTF is actually made up of three MTFs. First, there is the scanning distortion, which is due to the progression of the sea while the image is constructed by successive bands and leads to a rotation of the spectrum. This kind of distortion, however, is neglected in satellites, given their high speed and the matching rapidity of image construction. Another MTF is the tilt modulation, which is due to angle variations between the electromagnetic wave vector and the normal to the surface, again as a result of long waves. The third MTF is the hydrodynamic modulation that accounts for local variations in the small wave spectrum (especially as triggered by longer waves).

14.3.1. *Tilt modulation*

Let us imagine the surface of the sea as being divided into facets, or surface portions like the surface studied in section 14.2.2. Backscattering from these facets will then be due to ripple waves that are in turn modulated around a mean $E(\sigma^o) = I_0$ by long waves that change the angle between the incident wave vector and the normal (see Figure 14.3). Inside a resolution cell, i.e., the sea surface that a SAR pixel covers, every small portion of the surface, in which the small wave spectrum is generally constant, is a “facet” or a “backscatterer”. So backscattering for the entire resolution cell is the sum of all local backscatterings with some differences in electromagnetic wave progression that result in multiplicative noise.

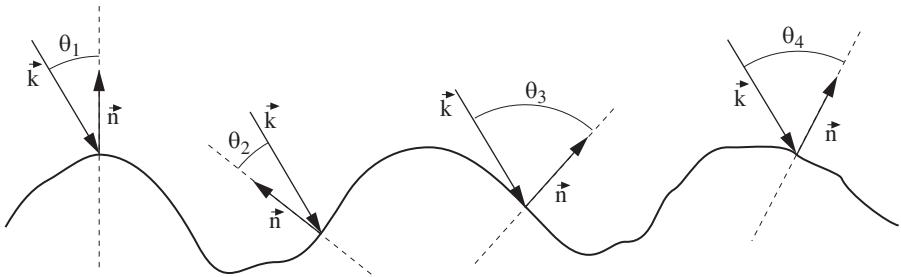


Figure 14.3. An example of electromagnetic reflection modulation by the sea surface

The backscattering section is a function of the angle θ between the incident wave vector and the normal. If we consider the facet slope as being defined by the angles α and φ , with α being the angle between the vertical and the projection of the normal on the incident plane, and φ being the angle between the vertical and the projection of the normal on the plane perpendicular to the incident plane, we have for the small angles [WRI 68]:

$$\sigma^{vv}(\theta_0, \alpha, \varphi) = g_{hh}(\theta_0 + \alpha)$$

$$\sigma^{vh}(\theta_0, \alpha, \varphi) = \tan(\varphi)[g_{vv}(\theta_0 + \alpha) - g_{hh}(\theta_0 + \alpha)]$$

$$\sigma^{vh}(\theta_0, \alpha, \varphi) = g_{hh}(\theta_0 + \alpha) + \frac{\tan^2(\varphi)}{\cos^2(\theta_0)} g_{vv}(\theta_0 + \alpha)$$

where θ_0 is the angle between the radius and a horizontal surface (modulation-free sea).

This model in which small waves modulated by long waves are responsible for the backscattering section is called the two-scale backscattering model, as it requires taking into account the spectra of both (i) the short waves that are involved in the backscattering section relation and the wavelengths of which vary from one centimeter to a dozen, depending on the considered band, and (ii) those of the long waves that intervene in modulation. Wright's results are in agreement with a great many experimental results concerning airborne radars [WRI 68]. It is worth noting that the rapid decrease of σ^{vv} and σ^{vh} (see Figure 14.4) for very small angles can

lead to non-linearities, since the first order linear approximation of σ^o (see equation [14.5]) does not hold in this area (especially not for σ^{vh}).

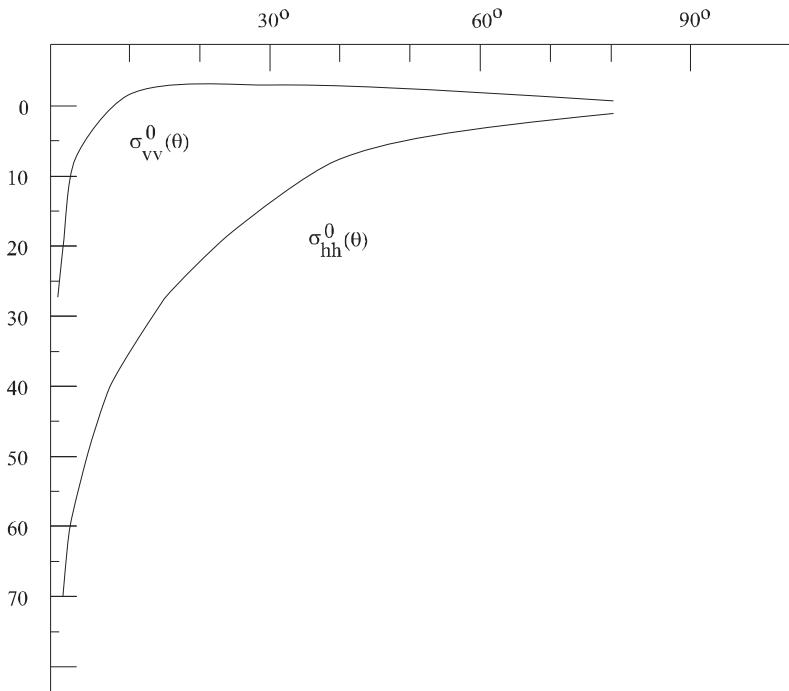


Figure 14.4. Backscattered section in terms of tilt angle, according to Wright [WRI 68]

This is the reason why the two-scale model cannot be fully used unless the angles range from 20° to 60° and ERS-1 having a tilt angle of 23° is at the limit of such model.

We know from section 14.2.1 that the slope and hence the angle between the normal and the incident radius is a linear function of the surface. In fact, the interesting relation here is that giving the backscattered section according to Fourier coefficients of the ocean rather than in terms of angle. To this end, it is generally more convenient to express the backscattered section in terms of its mean and local variation:

$$\sigma^o = I_0 + \delta\sigma^o = I_0 \left(1 + \int T^{tilt}(k) \zeta(k) e^{j\varphi(k)} e^{j(kx - \omega t_0)} dk \right) \quad [14.5]$$

where $T^{tilt}(k)$ is the tilt modulation transfer function, i.e., the transfer function between the surface and the backscattered modulation. We generally assume, as is implicitly done in the above expression, that the backscattering variation linearly depends on the angle between the normal and the incident wave vector and consequently on the frequency components of the sea. Keeping in mind the earlier formulated reservations, this is a generally acceptable assumption. The calculation of $T^{tilt}(k)$ (a Phillips spectrum (equation [14.4]) is assumed) leads to:

$$T^{tilt}(k) = 4jk_l \cot(\theta_0) \left(1 + \sin^2(\theta_0)\right)^{-1} \quad \text{for VV polarization}$$

$$T^{tilt}(k) = 8jk_l \left(\sin(\theta_0)\right)^{-1} \quad \text{for HH polarization}$$

where k_l is the projection of the wavenumber on the radar radius (under common conventions, k_l is identified with k_g). Figure 14.5 shows the module $T^{tilt}(k)$ for VV polarization. Note on both diagrams the linear dependence in terms of k_y .

However, these results are not the actually observed backscattering for two major reasons:

- as mentioned in the introduction, this model does not take into account the motion of the waves, which is a basic element in the backscattering of the facets, as will be shown again in section 14.3.2. This first approach, however, gives us an inkling that the small wave spectrum, to which backscattering is related, has no reason for being homogenous over time, given the motion of long waves and can therefore be modulated during surface illumination by the radar (hence the notion of the lifetime of a backscatterer [RAN 80]). Other SAR inherent phenomena further increase backscattering dependence on the long wave motion;

- the small wave spectrum does not necessarily have to be more homogenous in space than it is in time. Indeed, the long waves locally alter (or, as they say, “modulate”) the small wave spectrum. During his early work, Phillips found these phenomena that can be noticed even for a surface frozen in time, as will be shown at length in the next section.

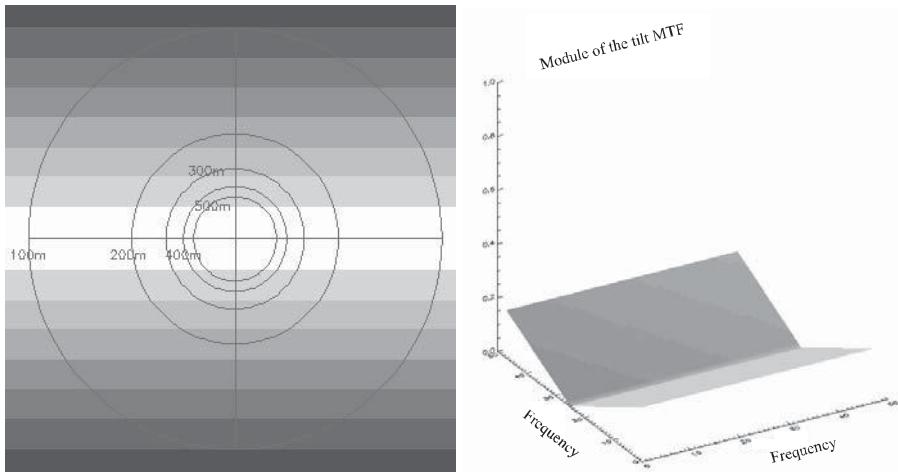


Figure 14.5. Module of the tilt modulation transfer function

14.3.2. Hydrodynamic modulation

The expression of the Phillips small wave spectrum that we used in the previous section to find the transfer function $T^{tilt}(k)$ actually illustrates an equilibrium mode, i.e., free of the modulation imposed by the long waves. Long wave acceleration and local velocity are responsible for the local variations of this spectrum. Keller and Wright [KEL 75] were the first ever to explain this phenomenon using the action “radiation” equations $N(x, k, t)$; for a sea wave with the wavenumber \hat{k} , the action and the spectrum are related by the expression:

$$N(x, \hat{k}, t) = S_D(x, \hat{k}, t) / \omega'$$

where ω' is the intrinsic pulse of the small waves, i.e., the pulse in the reference frame of long waves. In this section, we will denote by \hat{k} and $\hat{\omega}$, the values of the wavenumber and pulse, respectively, of small waves, in order to distinguish them from k and ω that refer to the long waves. Plant [PLA 86] and Alpers-Hasselmann [ALP 78] used the action directly and then deduced the variations of the spectrum based on action variations. To calculate small wave action scattering, Alpers-Hasselmann chose a moving reference frame related to the long waves. Within this frame, we can write the generic expression of action “radiation” as follows:

$$L[N(x, \hat{k}, t)] = \frac{dN(x, \hat{k}, t)}{dt} = \left(\frac{\partial}{\partial t} + \frac{\partial \omega}{\partial \hat{k}} \frac{\partial}{\partial x} - \frac{\partial \hat{\omega}}{\partial \hat{k}} \frac{\partial}{\partial \hat{k}} \right) N(x, \hat{k}, t) = Q(x, \hat{k}, t) \quad [14.6]$$

The source function $Q(x, \hat{k}, t)$ is in fact the sum of three different terms:

$$Q(x, \hat{k}, t) = Q_i(x, \hat{k}, t) + Q_r(x, \hat{k}, t) + Q_s(x, \hat{k}, t) \quad [14.7]$$

– $Q_i(x, \hat{k}, t)$ describes the energy transfer from the wind to the sea. Keller and Wright [KEL 75] as well as Plant [PLA 86] define this term as being proportional to action $Q_i(x, \hat{k}, t) = \beta \cdot N(x, \hat{k}, t)$;

- $Q_r(x, \hat{k}, t)$ accounts for energy transfer via wave interaction in the wave field;
- $Q_s(x, \hat{k}, t)$ draws together all the dissipation processes and accounts for energy losses.

The big problem with using the action “radiation” equation stems from the fact that there is very little understanding, if any, of these three source functions. As a result, some authors [KEL 75, ALP 78, PLA 86] have made several assumptions to help them use equation [14.6]. These authors have reached the following relation:

$$T^{hyd}(k) = -|k| \cdot \omega \cdot (4 + \gamma(\hat{k})) \cdot \left(\frac{k_y^2}{|k|^2} \right) \cdot \frac{\omega - j\mu}{\omega^2 + \mu^2}$$

They took into account that disturbances caused by long waves were weak around an equilibrium state $N_0(\hat{k})$ (i.e., the small wave spectrum undisturbed by long waves), which verified that $L_0[N_0(\hat{k})]$ and consequently $Q(x, \hat{k}, t)$ was zero.

The equilibrium state $N_0(\hat{k})$ simply translates the geometric effects related to the angle between the electromagnetic wave vector and the normal to the surface (what we are dealing with is the tilt modulation studied in the previous section).

Note, though, that while the measurements by [KEL 75] in a closed curve were in rather good agreement with theoretical results for low winds, this no longer applies when we have HH and VV polarization under high winds. Approximating the source function by a function proportional to $\delta N(x, \hat{k}, t)$ appears not to be appropriate any longer and the linear approximation of L may be called into question. Hasselmann and Hasselmann [HAS 91] have proposed a feedback term $Y_r + j \cdot Y_i$ (to be determined empirically) that helps ensure a better coherence of the results in the form:

$$T^{hyd}(k) = -|k|\omega(4 + \gamma(\hat{k})) \left[\frac{k_y^2}{|k|^2} + Y_r + jY_i \right] \cdot \frac{\omega - j\mu}{\omega^2 + \mu^2}$$

This MTF has an explicit input for the wind via the parameter μ . The module of the hydrodynamic transfer function is represented in Figure 14.6; the parameters here were selected based on $\mu(k, \hat{k}) = (0.04 \cdot u_{10} \frac{dk}{d\omega})^2 \hat{\omega}$, where u_{10} is the wind 10 m above sea level (in our case: 10 ms^{-1}) and $Y_r + jY_i = \mu \frac{dk}{d\omega} \exp(j\varphi)$ where $\varphi = \pi/4$.

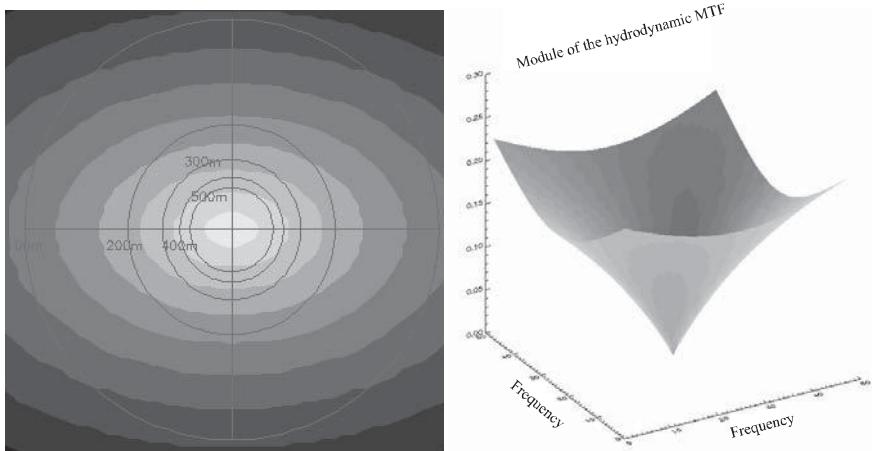


Figure 14.6. Module of the hydrodynamic modulation transfer function

Backscattering by the sea surface is thus made up of two modulations: one due to geometric effects (the tilt MTF for small wave equilibrium states) and a hydrodynamic modulation due to the effects of small wave modulation, i.e.:

$$T^{rar}(k) = T^{tilt}(k) + T^{hyd}(k).$$

As a result, the RAR pixel is expressed by the relation:

$$X_{rar}(x) = I_0 \cdot \left(1 + \int_{-\infty}^{+\infty} T^{rar}(k) \zeta(k) \cdot e^{j\varphi(k)} e^{j(kx - \omega t)} dk \right)$$

14.4. Sea surface motions

Object motion during radar illumination implies a phase change that hampers our estimate location on the image of the backscattering section of the signal. If the multiplicative coefficient of phase variation over time is V_s/R , the calculated displacement is $-u_r R t / V_s$, implying mislocation by a distance equal to:

$$d(x) = -\frac{u_r R}{V_s} \quad [14.8]$$

where R is the satellite to target distance, V_s is the satellite speed and u_r is the radial velocity of the reflector, as seen in section 9.1.5. For ERS-1, 1 ms^{-1} radial velocity implies a displacement by 115 m, or five to six pixels. In the sea surface case, this relation applies to the entire image, particularly as the sign of radial velocity is changed. The pixels corresponding to zero radial velocity are placed correctly, while the others are alternatively grouped together and dispersed according to the sign of radial velocity. This velocity bunching is shown in Figure 14.7.

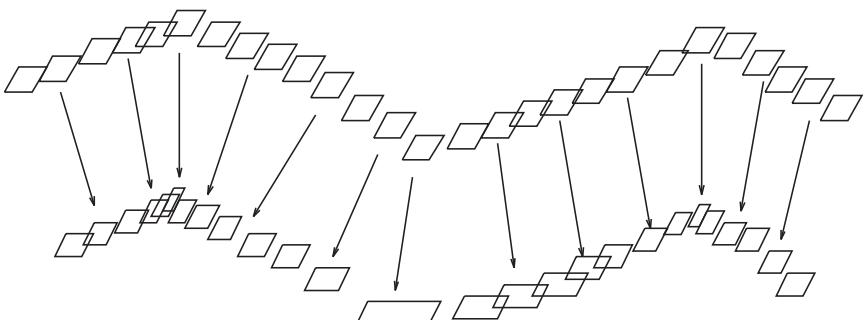


Figure 14.7. An example of SAR image creation from the surface of the sea; illustration of the phenomenon known as velocity bunching

Inside the resolution cell, these backscatterers have velocities almost equal to one another, but their dispersion leads to uneven displacements. The RAR pixel stretches or shrinks as a result and azimuth resolution is reduced. This azimuth smearing has long been discussed [RAN 80, TUC 83], but for the purpose of our work we will only retain the expression based on the SAR spectrum and mentioned in section 14.5.1.

14.5. SAR image of the sea surface

A coherent theory of the creation of the surface of the sea by SAR has taken a long time to become established, as the different views hampered the understanding of image creation by the synthetic aperture radars. The article by Hasselmann *et al.* [HAS 85] on the MARSEN experiment may be considered the first fruitful attempt at ironing out the differences and putting forward a unified theory. However, it was not until several years on that the expressions of the SAR spectrum (see section 14.5.1), the quasi-linear approximation and Volterra models (see section 14.5.2), and the inversion of the latter (see section 14.6.1) were given by Hasselmann and Hasselmann [HAS 91] and in a yet more elegant form by Krogstad [KRO 92].

14.5.1. Quasi-linear approximation of the SAR spectrum image

Assuming a linear transfer modulation, we know that the backscattered section is displaced by a distance $d(x)$. We also know that this section is a modulation by geometric and hydrodynamic phenomena around a mean. A SAR pixel is thus the sum of the RAR pixels displaced on its coordinates:

$$X_{SAR}(x) = \sum_{x'} X_{rar}(x') \left| \frac{dx'}{dx} \right|.$$

The sum is thus performed over the whole set of x' so that $x = x' - d(x')$. The Jacobian $\left| \frac{dx'}{dx} \right|$ is the expression of the size variation of the backscattered element due to the velocity dispersion phenomenon discussed above. To calculate the theoretical SAR spectrum and bispectrum, we have to assume sea surface stationarity (up to the third order at least). Sea surface ergodicity needs to be assumed in order to estimate the moments and the spectra by averaging over the spatial domain. In a first stage, the SAR process mean can be expressed as:

$$\begin{aligned}
E(X_{SAR}) &= \lim_{|A| \rightarrow +\infty} |A|^{-1} \int_A X_{SAR}(x) dx \\
&= \lim_{|A| \rightarrow +\infty} |A|^{-1} \int_A \sum_{x'} X_{rar}(x') \left| \frac{dx'}{dx} \right| dx \\
&= \lim_{|A| \rightarrow +\infty} |A|^{-1} \int_{A'} X_{rar}(x') dx' = E(X_{rar}) = I_o
\end{aligned}$$

The SAR process thus does not cause any variation in the backscattered section mean. The Fourier coefficients are given by:

$$\begin{aligned}
\tilde{X}_{SAR}(k) &= |A|^{-1} \int_A X_{SAR}(x) e^{-jkx} dx \\
&= |A|^{-1} \int_A \sum_{x'} X_{rar}(x') \left| \frac{dx'}{dx} \right| e^{-jkx} dx \\
&= |A|^{-1} \int_{A'} X_{rar}(x') e^{jk_x d(x')} e^{-jkx'} dx'
\end{aligned} \tag{14.9}$$

Therefore, it is the exponential term in $Y(x') = X_{rar}(x') e^{jk_x d(x')}$ that contains non-linearity. The spectrum expression is given by [KRO 92]:

$$S_{SAR}(k) = I_0^2 e^{-k_x^2 M_2^{dd}(0)} \int_{-\infty}^{+\infty} e^{k_x^2 M_2^{dd}(x)} G(x, k_x) e^{-jk(x)} dx - I_0^2 \delta(k)$$

where:

$$\begin{aligned}
G(x, k_x) &= 1 + M_2^{rr}(x) + jk_x (M_2^{rd}(x) + M_2^{rd}(-x)) \\
&\quad + k_x^2 (M_2^{rd}(0) - M_2^{rd}(x)) (M_2^{rd}(0) - M_2^{rd}(-x))
\end{aligned}$$

where $M_2^{dd}(x) = E(d(x')d(x'+x))$ is the displacement autocorrelation function; $M_2^{rr}(x) = E(X_{rar}(x')X_{rar}(x'+x))$ is the autocorrelation function of the RAR pixels and $M_2^{rd}(x) = E(X_{rar}(x')d(x'+x))$ is the correlation function between

RAR pixels and displacements. The term $e^{-k_x^2 M_2^{dd}(0)}$ is the expression of azimuth smearing, the effects of which tend to filter out high azimuthal components, thus decreasing resolution in this direction.

In order to show the effects of non-linearity, we need to calculate the bispectrum of the image, i.e., the expression $E\left(\tilde{X}_{SAR}(k_1)\tilde{X}_{SAR}(k_2)\tilde{X}_{SAR}^*(k_1+k_2)\right)$ [LEC 02].

The expressions of the spectrum (particularly those of the bispectrum) are hard to interpret. Note that radial velocity is a major non-linearity factor. The higher this velocity, the more significant the SAR image distortion (owing to the displacement content of the exponential). It is quite evident that the higher the sea surface elevation (and hence the significant height), the higher the velocity. However, the higher the velocity, the more significant the azimuth smearing is too. Azimuth filtering and energy creation via non-linear interactions are competing phenomena.

If we assume that k_x is small and the velocities are low enough for the exponential to be approachable by $1+k_x^2 M_2^{dd}(x)$, we obtain:

$$\begin{aligned} S_{SAR}(k) &\approx I_0^2 \int_{-\infty}^{+\infty} \left(1 + k_x^2 \cdot M_2^{dd}(x) + M_2^{rr}(x)\right) e^{-jkx} dx - I_0^2 \delta(k) \\ &\approx k_x^2 \cdot S_{dd}(k) + S_{rar}(k) \end{aligned}$$

The SAR image spectrum consequently is made up of two parts: one is related to the sea surface motion and another to the geometric and hydrodynamic backscattering of the surface (RAR modulation).

The modulation transfer function of velocity bunching is then expressed by taking into account the exponential approximation and equation [14.2] as:

$$T^{vb}(k) = j \cdot k_x \cdot d(x') = -\frac{R}{V_s} k_x \sqrt{gk} \left(\cos \theta - j \frac{k_y}{k} \sin \theta \right).$$

The SAR spectrum is obtained from:

$$S_{SAR}(k) = |T^{SAR}(k)|^2 S_{sea}(k) \quad \text{with} \quad T^{SAR}(k) = T^{vb}(k) + T^{rar}(k).$$

This approximation in which the SAR spectrum and sea spectrum are linearly related is called a quasi-linear approximation. Since the velocity bunching modulation is stronger than the RAR modulation, it is the latter that will impose its form on the SAR spectrum and since it has a zero on the radial axis, a spectrum splitting (see Figure 14.9) can be seen when the spectrum itself lies on that axis. Moreover, in the case of a quasi-linear phenomenon, the SAR image bispectrum is zero, as it is the bispectrum of a linear filter output that has a Gaussian data input.

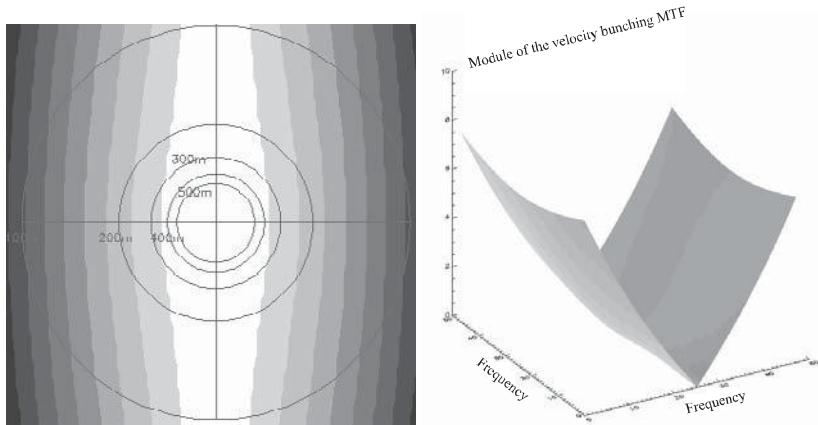


Figure 14.8. Module of the velocity bunching modulation transfer function

14.5.2. Modeling of non-linearities

In order to understand SAR non-linearities in spectral and bispectral terms, Volterra models are used to decompose the SAR transform, as they help obtain a decomposition, part of which is linear while another part is strictly non-linear. In the frequency domain, these models are written:

$$\tilde{Y}(k) = \sum_{n=0}^{+\infty} \tilde{X}_n(k)$$

where

$$\begin{aligned} \tilde{X}_n(k) &= \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} H_n(k - k_1 - \dots - k_{n-1}, k_1, \dots, k_{n-1}) \\ &\quad \tilde{X}(k - k_1 - \dots - k_{n-1}) \tilde{X}(k_1) \dots \tilde{X}(k_{n-1}) dk_1 \dots dk_{n-1} \end{aligned}$$

The SAR process decomposition [LEC 02] (or in other words, the expansion of the exponential) makes it possible to find the $H_n(k)$ kernels. The SAR transform quadratic kernel, for example, is:

$$H_2(k_1, k - k_1) = \left[\frac{1}{2} \left(jk_x \frac{R}{V} \right)^2 D(k_1) D(k - k_1) + \left(jk_x \frac{R}{V} \right) D(k_1) T^{rar}(k - k_1) \right] \tilde{X}(k_1) \tilde{X}(k - k_1) \quad [14.10]$$

The product $\tilde{X}(k - k_1) \tilde{X}(k_1)$ shows that in the case of this kernel, a wave with a wave vector $k - k_1$ interacts quadratically with a wave having a wave vector k_1 to generate a wave with a wave vector k , which is an artifact of the SAR mapping. To put this better, the SAR process generates “waves” that are not geophysical information.

The results of the Volterra model show that the SAR process for a common sea state ($H_s \leq 7$ m and less than 200 m wavelength) is dominated by a linear process with cubic interactions that may be non-negligible (see Figure 14.9). In extreme sea conditions or azimuth propagation cases (see Figure 14.10), the process becomes rapidly dominated by quadratic interactions (and even order, up to sixth order). This rapid change arising in such severe sea states stems on the one hand from an increased number of quadratic interactions (situated within a band along the radial axis) and on the other hand, from the azimuth smearing that filters out the linear/cubic components.

In every scenario (whether linear or not), velocity bunching, which is far superior to the RAR modulation, is the prevailing mechanism. Finally, non-linearity analysis, based on higher order spectra and the Volterra model, indicates that non-linearities are easier to detect, locate and quantify in weakly non-linear images than in strongly non-linear images. In this case, the numerous input pixels convoluted with non-linear kernels make us return to Gaussian statistics (central limit theorem) and hence the bispectrum is zero. The Volterra models theoretically allow a non-linear inversion in some conditions that are not fully verified for the SAR transform as a result of azimuth defocusing. In the next section, we will confine ourselves to presenting the traditional theory of inversion of the long wave SAR spectrum

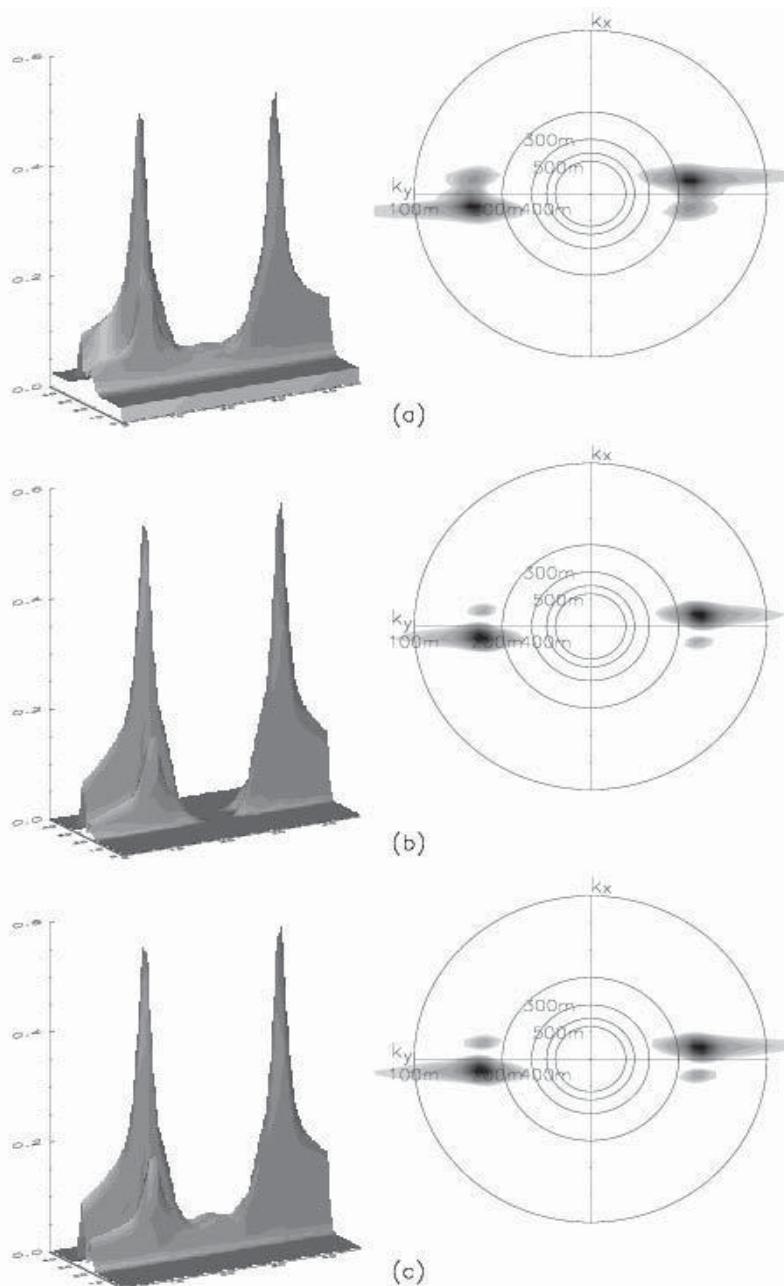


Figure 14.9. SAR spectrum (a), first order (linear, b) and second order Volterra model spectrum for a radially propagating swell ($H_S = 7 \text{ m}$ and wavelength = 200 m)

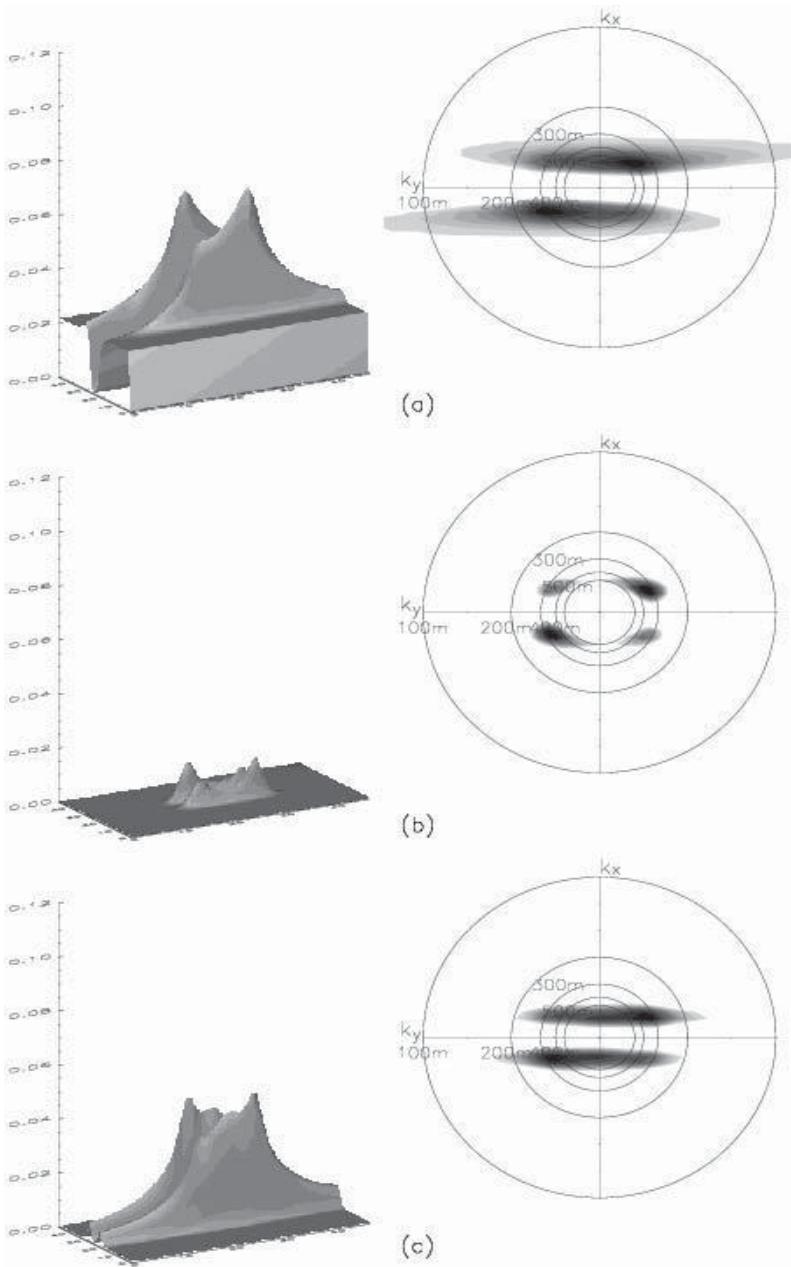


Figure 14.10. SAR spectrum (a), first order (linear, b) and second order Volterra model spectrum for a swell propagating in the azimuth ($H_s = 7\text{ m}$ and wavelength = 200 m)

14.6. Inversion of the SAR imaging mechanism

Two inversion methods, one related to the long wave spectrum and another to detect and measure seabed bathymetry by radar imaging, are among the chief applications of the approaches described earlier. The principles of these methods are detailed in the following sections that focus on the uses of “system” modeling (see section 14.5.2) by taking into account the intrinsic non-linearities resulting from SAR image acquisition. As mentioned in our introduction to this chapter, the ocean surface images produced by SAR have boosted the development of some analysis techniques aimed at improving the characterization of surface phenomena that modulate the roughness to which radars are sensitive.

14.6.1. SAR spectrum inversion

“Inverting” a SAR spectrum consists of retrieving the sea spectrum from a SAR spectrum. In other words, it implies restoring a component removed by the azimuth smearing and removing non-linear artifacts. The only method to do this has been proposed by Hasselmann and Hasselmann [HAS 91] and refined in [HAS 96]. Mastenbroek introduced some important improvements using external ocean parameters. Finally, Lyzenga proposed a self-sufficient algorithm (by extracting these parameters from the SAR image).

This makes it possible to estimate the sea spectrum iteratively by minimizing a regularization cost function. For every iteration, the cost function is defined by:

$$J = \sum_k \left[S_{SAR}^{(n)}(k) - \hat{S}_{SAR}(k) \right]^2 + \mu \left[\frac{S_{sea}^{(n)}(k) - S_{sea}^{fg}(k)}{B + S_{sea}^{fg}(k)} \right]^2$$

where $S_{sea}^{(n)}(k)$ is the sea spectrum in the n^{th} iteration, $S_{SAR}^{(n)}(k)$ is the SAR spectrum corresponding to this sea spectrum, $S_{sea}^{fg}(k)$ is the first guess spectrum (that final results heavily depend on) and $\hat{S}_{SAR}(k)$ is the observed SAR spectrum. To estimate the sea spectrum in the $n+1^{\text{th}}$ iteration, we add a spectrum variation $S_{sea}^{(n)+1}(k) = S_{sea}^{(n)}(k) + \Delta S_{sea}^{(n)}(k)$. The term $\Delta S_{sea}^{(n)}(k)$ is proportional to the SAR spectrum variation $\Delta S_{SAR}^{(n)}(k)$.

As $S_{sea}^{(n)+1}(k)$ is calculated, the new SAR spectrum is obtained by quasi-linear transform. The algorithm is stopped when the values of $\Delta S_{SAR}^{(n)}(k)$ are small enough. This inversion method leads to good results unless azimuth smearing has “turned” the spectrum, as sometimes happens. The reader must keep in mind that the “SAR spectrum” assumes a Gaussian sea surface and then is only valid for open seas.

14.6.2. Underwater topography inversion

The modeling process is here applied to a full imaging chain, i.e., from the seabed-surface-current transfer function (action equation) to the SAR transform. This model (in the sense of the Volterra model described in section 14.5.2) may then, under certain conditions, be reversed thus yielding a measure of the digital terrain model (DTM) associated with the underwater relief.

The imaging mechanism of the underwater radar topography, as is widely admitted nowadays, is comprised of three stages (see Figure 14.11):

- the relief modulates the current moving over it;
- the waves obtain energy from the current;
- the radar can capture the energy backscattered from the waves having a certain wavelength.

While the topography-to-current transfer can be modeled by a simple continuity equation, the two other stages require very complex physical modeling.

Energy transfer from currents to the small waves is described by the conservation energy introduced earlier in equation [14.6]. More specifically, this equation relates the current $U(x)$ to the action $N(x, \hat{k}, t)$. Taking into account bathymetry-induced currents, the action equation takes up the following form [ROM 97]:

$$\begin{aligned} \frac{\partial N(x, \hat{k}, t)}{\partial t} + (c_g(k) + U(x)) \frac{\partial N(x, \hat{k}, t)}{\partial x} \\ - k \frac{\partial U(x)}{\partial x} \frac{\partial N(x, \hat{k}, t)}{\partial k} = Q(x, \hat{k}, t) \end{aligned} \quad [14.11]$$

where $Q(x, \hat{k}, t)$ is the source term that turns the disturbed spectrum back to balance. Action modulation with respect to its equilibrium is here driven by the current $U(x)$. By introducing $N(x, \hat{k})$ in the action conservation equation, we obtain Volterra kernels of the type shown in equation [14.10]. To calculate these kernels, we first need to identify the degrees of interactions (linear, quadratic, etc.) that are involved. The Fourier coefficients of the SAR image are expressed by equation [14.9] and thus bring into play non-linearities related to the exponential function. The Volterra model that we can also find by expanding the exponential makes it possible to obtain the RAR pixel based on the surface current vector. The inversion algorithm of the image in current values can be divided into two distinct parts. The first part refers to the regularization of observed data and the second to mechanism inversion for the ideal case of an error-free observation. Once the surface current has been obtained, depth is estimated using the continuity equation. Figure 14.12 [ING 99b] illustrates the comparison between a real DTM and that obtained by the inversion method described above.

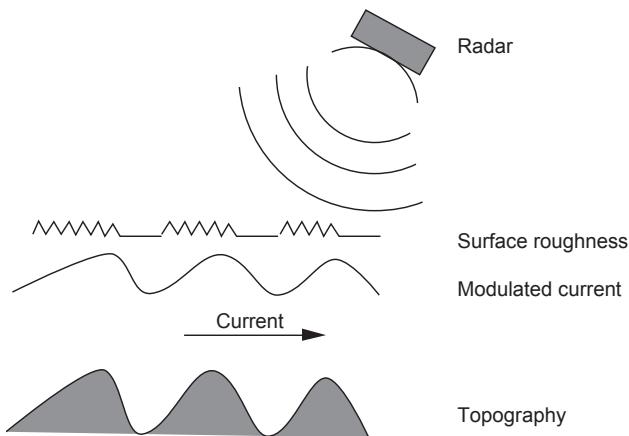


Figure 14.11. Imaging mechanism diagram

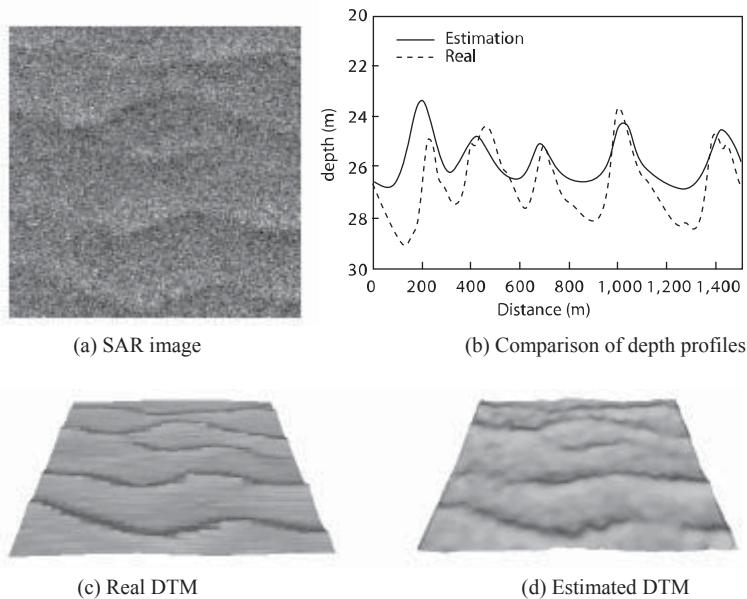


Figure 14.12. Inversion by Tikhonov regularization of the image CSTAR-105-3

This page intentionally left blank

Bibliography

- [ABD 98] ABDELFATTAH R., NICOLAS J., BOUSSEMA M. “Topography Optimisation for Accurate Interferometric Phase Generation”. *EUROPTO*, 251–256, Barcelona, Spain, 1998.
- [AIN 00] AINSWORTH T., LEE J., SCHULER D. “Multi-frequency polarimetric SAR data analysis of Ocean surface Features”. *3rd European Conference on Synthetic Aperture Radar, EUSAR 2000*, 1, 281–284, Berlin, 2000. VDE Verlag.
- [ALP 78] ALPERS W., HASSELMANN K. “The Two Frequency Microwave Technique for Measuring Ocean-Wave Spectra from an Airplane or Satellite”. *Boundary Layer Meteorol.*, 13:215–230, 1978.
- [ANO 93] ANONYMOUS. “Special Issue on RADARSAT”. *Canadian Journal of Remote Sensing*, 1993.
- [ARS 84] ARSENAULT H. H., LEVESQUE M. “Combined homomorphic and local statistics processing for restoration of images degraded by signal-dependent noise”. *Applied Optics*, 23(6):845–850, 1984.
- [AYA 89] AYACHE N. *Vision stéréoscopique et perception multisensorielle*. InterEditions, Paris, 1989.
- [BEA 66] BEAN B., CAHOON B., SAMSON C., THAYER G. *A World Atlas of Atmospheric Radio Refractivity*. US Government Printing Office. ESSAMonograph 1, 1966.
- [BEC 63] BECKMANN P., SPIZZICHINO A. *The Scattering of Electromagnetic Waves from Rough Surfaces*. Macmillan, New York, 1963.
- [BEC 87] BECKMAN P., SPIZZICHINO A. *The Scattering of Electromagnetics Waves from Rough Surfaces*. Artech House, Norwood MA, 1987.
- [BES 86] BESAG J. “On the statistical analysis of dirty pictures”. *JRSS B*, 48:259–302, 1986.
- [BIA 99] BIANCHIN A., MAÎTRE H., NICOLAS J.-M., TUPIN F. “Monitoring urban areas (Paris) with multitemporal multincidence ERS images”. *CEOS99*, Toulouse, France, 1999.

- [BIJ 97] BIJAOUI A., ROBICHON Y., FANG Y., RUÉ F. "Méthodes multi-échelles appliquées à l'analyse des images radar à ouverture synthétique". *Traitements du Signal*, 14(2):179–193, 1997.
- [BLA 93] BLACKNELL D., OLIVER C. "Comments on Statistical and spatial properties of forest clutter measured with polarimetric SAR". *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):1093–1094, 1993.
- [BLA 94] BLACKNELL D. "Comparison of parameter estimators for K-distribution". *IEEE Proc. Radar, Sonar Navigation*, 141(1):45–52, 1994.
- [BON 91] BONE D. "Fourier fringe analysis: the two-dimensional phase unwrapping problem". *Applied Optics*, 30(25):3627–3632, 1991.
- [BOR 65] BORN M., WOLF E. *Principles of Optics*. Pergamon Press, New York, 1965.
- [BOR 87] BORGEAUD M., SHIN R., KONG J. "Theoretical models for polarimetric radar clutter". *Journal of Electromagnetic Waves and Applications*, 1(1):73–79, 1987.
- [BOV 88] BOVIK A. C. "On Detecting Edges in Speckle Imagery". *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-36(10):1618–1627, 1988.
- [BRU 88] BRUCKSTEIN A. "On shape from shading". *Computer Vision, Graphics and Image Processing*, 44(2):139–154, 1988.
- [BRU 94] BRUNIQUEL J., LOPÈS A. "Analysis and Enhancement of multitemporal SAR data". *Image and Signal Processing of Remote Sensing*, 2315:342–353, 1994.
- [BRU 96] BRUNIQUEL J. "Contribution de données multi-temporelles à l'amélioration radiométrique et à l'utilisation d'images de radar à synthèse d'ouverture". 2245, Thesis, Paul Sabatier University (Toulouse), 1996.
- [BRU 97] BRUNIQUEL J., LOPÈS A. "Multivariate speckle reduction in SAR images". *International Journal of Remote Sensing*, 18(3):603–627, 1997.
- [BRU 98] BRUNIQUEL J., LOPÈS A. "True Multilook Intensity Distribution in SAR Data Imagery". *IGARSS'98*, A01-02.pdf, Seattle, 1998.
- [CAR 91] CARIOU C. *Analyse spectrale d'images numériques. Application à la télédétection spatiale des océans*. ENST Bretagne – Thesis from the University of West Brittany, 1991.
- [CAR 94] CARAMMA M. "Analyse Spectrale: application de l'algorithme MUSIC à la caractérisation de franges en interférométrie SAR". Technical report ENST, 1994.
- [CAV 93] CAVES R. G. "Automatic matching of features in Synthetic Aperture Radar data to digital map data". PhD Thesis, University of Sheffield, 1993.
- [CAV 98] CAVES R., QUEEGAN S., WHITE R. "Quantitative comparison of the performance of SAR segmentation algorithms". *IEEE Transactions on Geoscience and Remote Sensing*, 7(11):1534–1546, 1998.
- [CHA 89] CHALMOND B. "An iterative Gibbsian technique for the reconstruction of M-ary image". *Pattern Recognition*, 22(6):747–761, 1989.
- [CHA 94] CHAPMAN B. "SIR-C data compression software user guide". D-11427, JPL Pasadena USA, 1994.

- [CHA 99] CHANUSSOT J., MAURIS G., LAMBERT P. "Fuzzy Fusion Techniques for linear features detection in multi-temporal SAR images". *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1292–1305, 1999.
- [CHE 96] CHEN K. S., HUANG W. P., TSAY D. H., AMAR F. "Classification of Multifrequency Polarimetric SAR Imagery Using a Dynamic Learning Neural Network". *IEEE Transactions on Geoscience and Remote Sensing*, 34(3):814–820, 1996.
- [CIL 97] CILF. *Terminologie de télédétection et photogrammétrie*. Conseil International de la Langue Française – PUF, 1997.
- [CLO 95] CLOUDE S. "Symmetry, zero correlations and target decomposition theorems". *Proceedings of the 3rd International Workshop on Radar Polarimetry, JIPR'95*, IRESTE, University of Nantes, 1995.
- [CLO 96] CLOUDE S., POTTIER E. "An entropy based classification scheme for land applications of polarimetric SAR". *IEEE Transactions on Geoscience and Remote Sensing*, 35(1):68–78, 1996.
- [COC 96] COCQUEREZ J., PHILIPP S. *Analyse d'Images: Filtrage et Segmentation*. Masson, 1996.
- [COH 60] COHEN J. "A review of assessing the accuracy of classification of remotely sensed data". *Remote Sensing of Environment*, 37:35–46, 1960.
- [CON 87] CONTE E., LONGO M. "Characterization of radar clutter as a spherically invariant random vector". *IEEE proc. F. Commun. Radar Signal Processing*, 134:191–197, 1987.
- [CRI 85] CRIMMINS T. "Geometric filter for speckle reduction". *Applied Optics*, 24(10):1438–1443, 1985.
- [CRI 86] CRIMMINS T. "Geometric filter for reducing speckle". *Optical Engineering*, 25(5):651–654, 1986.
- [DAB 95] DABA J., BELL M. R. "Statistics of the scattering cross-section of a small number of random scatterers". *IEEE Trans. on Antennas and Propagation*, 43(8):773–783, 1995.
- [DAI 84] DAINTY J. *Laser Speckle and Related Phenomena*. Springer-Verlag, 1984.
- [DAN 86] DANA R., KNEPP D. "The impact of strong scintillation on space based radar design II: noncoherent detection". *IEEE Trans. on Aerospace and Electronic Systems*, 22(1):34–46, 1986.
- [DAT 98] DATCU M., SEIDEL K., WALESSA M. "Spatial Information Retrieval from Remore-Sensing Images. Part I: Information Theoretical Perspective". *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1431–1445, 1998.
- [DEL 93] DELIGNON Y. "Etude statistique d'images radar de la surface de la mer". PhD Thesis, University of Rennes 1, 1019, 1993.
- [DEL 97a] DELIGNON Y., GARELLO R., HILLION A. "Statistical Modelling of Ocean SAR Images". *IEEE Proc. Radar Sonar and Navigation*, 144(6):348–354, 1997.

- [DEL 97b] DELIGNON Y., MARZOUKI A., PIECZYNSKI W. "Estimation of generalised mixture and unsupervised statistical radar image segmentation". *IEEE Trans. on Image Processing*, 6(10):1364–1375, 1997.
- [DEM 77] DEMPSTER N., LAIRD N., RUBIN D. "Maximum likelihood from incomplete data via the EM algorithm". *Journal of the Royal Statistical Society, B*, 9:1–38, 1977.
- [DER 87] DERIN H., ELLIOTT H. "Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-9(1):39–55, 1987.
- [DER 90] DERIN H., KELLY P. A., VEZINA G., LABITT S. G. "Modeling and Segmentation of speckled Images Using Complex Data". *IEEE Transactions on Geoscience and Remote Sensing*, 28(1):76–87, 1990.
- [DER 95] DERAUW D. "Phase unwrapping using coherence measurements". *Synthetic Aperture Radar and Passive Microwave Sensing, Paris*, 2584 EUROPTO, 319–324. SPIE, 1995.
- [DES 95] DESCOMBES X. "A fission and fusion Markovian approach for multi-channel segmentation". *IGARSS'95*, Florence, Italy, 1, 124–126, 1995.
- [DON 98] DONG Y., FORSTER B. C., MILNE A. K., MORGAN G. A. "Speckle suppression using recursive wavelet transforms". *International Journal of Remote Sensing*, 19(2):317–330, 1998.
- [DUB 90] DUBUISSON B. *Diagnostic et reconnaissance des formes*. Hermès, Paris, 1990.
- [DUB 95] DUBOIS P., VAN ZYL J., ENGMAN T. "Measuring soil moisture with imaging radars". *IEEE Transactions on Geoscience and Remote Sensing*, 33:915–926, 1995.
- [DUP 95] DUPONT S., LABROUSSE D., BERTHOD M. "SAR interferometry: an application of simulated annealing to phase unwrapping". *9th Scandinavian Conference on Image Analysis, Uppsala, Sweden*, 697–705, 1995.
- [DUP 97] DUPONT S., NONIN P., RENOUARD L., PICHON G., BIGNONE F. "Atmospheric artifacts in ERS DTM. ISTAR experience over multiple sites and large areas". *3rd ERS Symposium on Space at the Service of our Environment*, 414, 1711–1719, Florence (Italy), March 1997. ESA-SP.
- [DUP 98] DUPONT S., DEMARGNE L., DURAND P. "Production of digital elevation models by radargrammetry using RadarSat data". *Final Adro Symposium*, Montreal (Canada), 1998.
- [DUR 87] DURAND J. M. "Theoretical feasibility of SAR interferometry". *Proc IGARSS'87*, Ann-Harbor (Michigan), 1987.
- [ELI 00] ELIZAVETIN I. V. "Approach to the SAR stereo data processing based on the matching of multiangle set". *EUSAR, 3rd European Conference on Synthetic Aperture Radar*, 1, 427–430, Berlin, 2000. VDE Verlag.
- [FIS 81] FISCHLER M. A., TENENBAUM J., WOLF H. "Detection of roads and linear structures in low resolution aerial imagery using a multi-source knowledge interpretation technique". *Computer Graphics and Image Processing*, 15:201–223, 1981.

- [FJO 97] FJORTOFT R., MARTHON P., LOPÈS A., CUBERO-CASTAN E. "Multi-edge detection in SAR images". *ICASSP'97*, 1997.
- [FJO 98] FJORTOFT R., LOPÈS A., MARTHON P., CUBERO-CASTAN E. "An Optimal Multiedge Detector for SAR Image Segmentation". *IEEE Transactions on Geoscience and Remote Sensing*, 36(3):793–802, 1998.
- [FJO 99a] FJORTOFT R. "Segmentation d'images radar par détection de contours". No. 1535, Inst. Nat. Polytechnique de Toulouse 1999.
- [FJO 99b] FJORTOFT R., LOPÈS A., BRUNIQUEL J., MARTHON P. "Optimal edge detection and edge localization in complex SAR images with correlated speckle". *IEEE Trans. Geoscience and Remote Sensing*, 37(5):2272–2281, 1999.
- [FOR 95] FORNARO G., FRANCESCHETTI G. "Image registrationin interferometric SAR processing". *IEE Proc. Radar, Sonar Navig*, 142:313, 1995.
- [FOR 96] FORNARO G., FRANCESCHETTI G., LANARI R. "Interferometric phase unwrapping using Green's formulation". *IEEE Trans. Geosci. and Remote Sensing*, 34:720–727, 1996.
- [FOR 97] FORNARO G., FRANCESCHETTI G., LANARI R., SANSOSTI E., TESAURO M. "Global and local phase-unwrapping techniques: a comparison". *Journal of the Optical Society of America A*, 14:2702–2708, 1997.
- [FOU 96] FOUCHER S., BÉNIÉ G. B., BOUCHER J.-M. "Unsupervised multiscale speckle filtering". *ICIP-96*, I:391–394, 1996.
- [FOU 99] FOUCHER S., ZAGOLSKY F., GAILLARD C., VOIRIN Y., NGUYEN M. N. "Influence of speckle filtering on the assessment of agricultural surface parameters". *IGARSS'99*, Hamburg, Germany, 1999.
- [FRA 70] FRANÇON M. *Vibrations lumineuses. Optique cohérente*. Dunod, 1970.
- [FRA 87] FRANKOT R., CHELLAPPA R. "Application of a Shape from Shading Technique to Synthetic Aperture Radar Imagery". *Proceedings of IGARSS'87 Symposium*, 1323–1329, 1987.
- [FRA 90] FRANKOT R. CHELLAPPA R. "Estimation of surface topography from SAR imagery using shape from shading algorithms". *Artificial Intelligence*, 43:271–310, 1990.
- [FRA 95] FRANCESCHETTI G., PASCAZIO V., SCHIRINZI G. "Iterative homomorphic technique for speckle reduction in SAR imaging". *J. Opt. Soc. Am. A*, 12(4):686–694, 1995.
- [FRE 92a] FREEMAN A. DURDEN S. "A three component scattering model to describe polarimetric SAR data". *Radar Polarimetry, SPIE*, 1748, 213–225, 1992.
- [FRE 92b] FREEMAN A., VAN ZYL J., KLEIN J., ZEBKER H., SHEN Y. "Calibration of Stokes and scattering matrix format polarimetric SAR data". *IEEE Transactions on Geoscience and Remote Sensing*, 30:531–539, 1992.
- [FRE 94] FREEMAN A., VILLASENOR J., KLEIN J., HOOGEBOOM P., GROOT J. "On the use of multi-frequency and polarimetric radar backscatter features for classification of agricultural crops". *International Journal of Remote Sensing*, 15(9):1799–1812, 1994.

- [FRE 97] FRERY A., MULLER H., YANASSE C., SANT'ANNA S. "A model for extremely heterogeneous clutter". *IEEE Trans GRS*, 35(3):648–659, 1997.
- [FRO 82] FROST V. S., STILES J. A., SHANMUGAN K. S., HOLTZMAN J. C. "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):157–166, 1982.
- [FRO 85] FROST V., YUROVSKY L. "Maximum likelihood classification of Synthetic Aperture Radar Imagery". *Comp. Vision Graphics and Image Processing*, 32(3):291–313, 1985.
- [FU 76] FU K. *Digital Pattern Recognition*. Springer-Verlag, New York, 1976.
- [FUN 94] FUNG A. *Microwave Scattering and Emission Models and their Applications*. Artech House Inc., Norwood (MA-USA), 1994.
- [GAB 88] GABRIEL A. K., GOLDSTEIN R. M. "Crossed orbit interferometry: theory and experimental results from SIR-B". *Int. J. Remote Sensing*, 9(5):857–872, 1988.
- [GAR 07] GARELLO R. *Two-dimensional Signal Analysis*, ISTE Ltd, London, forthcoming.
- [GAT 94] GATELLI F., GUARNIERI A. M., PARIZZI F., PASQUALI P., PRATI C., ROCCA F. "The wavenumber shift in SAR interferometry". *IEEE Trans. Geosci. and Remote Sensing*, 32(4):855–865, 1994.
- [GEM 84] GEMAN S., GEMAN D. "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of Images". *IEEE-PAMI*, 6:721–741, 1984.
- [GER 98] GERMAIN O., RÉFRÉGIER P. "Edge localisation and detection in SAR images: a comparative study of global filtering and active contour approaches". *EUROPTO '98*, Barcelona, Spain, 3500:111–121, 1998.
- [GHI 87] GHIGLIA D. C., MASTIN G. A., ROMERO L. A. "Cellular-automata method for phase unwrapping". *J. Opt. Soc. Am. A*, 4(1):267–280, 1987.
- [GHI 89] GHIGLIA D. C., ROMERO L. A. "Direct phase estimation from phase differences using fast elliptic partial differential equation solvers". *Optical Letters*, 14:1107–1109, 1989.
- [GHI 94] GHIGLIA D. C. and ROMERO L. A. "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods". *J. Opt. Soc. Am. A*, 11(1):107–117, 1994.
- [GHI 98] GHIGLIA D. C., PRITT M. D. *Two-Dimensional Phase Unwrapping*. John Wiley and Sons, New York, 1998.
- [GIE 87] GIERLOFF J. J. "Phase unwrapping by regions". *Current Developments in Optical Engineering II*, 818, 2–9. SPIE, 1987.
- [GOL 88] GOLDSTEIN R. M., ZEBKER H. A., WERNER C. L. "Satellite radar interferometry: two-dimensional phase unwrapping". *Radio Science*, 23:713–720, 1988.
- [GOL 89] GOLUB G. H., LOAN C. F. V. *Matrix Computations*, Chapter 10, 505–538. The Johns Hopkins University Press, London, 1989.

- [GOO 63] GOODMAN N. "Statistical analysis based on a certain multivariate complex Gaussian distribution". *Ann. Math. Statist.*, 34:152–177, 1963.
- [GOO 72] GOODMAN J. *Introduction à l'optique de Fourier et à l'holographie*. Masson, 1972.
- [GOO 75] GOODMAN J. "Statistical properties of laser speckle patterns". *Laser Speckle and Related Phenomena*, 9, 9–75. J.C. Dainty (Springer-Verlag, Heidelberg, 1975), 1975.
- [GOO 76a] GOODMAN J. "Some fundamental properties of speckle". *Journal Optical Society of America*, 66(11):1145–1150, 1976.
- [GOO 76b] GOODMAN J. *Statistical Optics*. John Wiley, New York, 1976.
- [GOO 86] GOODMAN J. "A random walk through the field of speckle". *Optical Engineering*, 25(5):610–612, 1986.
- [GOU 96] GOUINAUD C. "Traitement d'images satellitaires pour la detection d'agglomérations". PhD Thesis, Télécom Paris ENST 96E035, 1996.
- [GRA 97] GRASSIN S. *Analyse et interprétation spatiale d'images SAR de la surface de l'océan*. ENST Bretagne – Thesis from the University of Rennes I, 1997.
- [GUI 90] GUINDON B. "Development of a shape from shading technique for the extraction of topographic models from individual spaceborne SAR images". *IEEE Trans on Geoscience and Remote Sensing*, 28(4):654–661, 1990.
- [HAR 94] HARA Y., ATKINS R., YUEH S., SHIN R., KONG J. "Application of neural networks to radar image classification". *IEEE Transactions on Geoscience and Remote Sensing*, 32:100–109, 1994.
- [HAS 61] HASSELMANN K. "On the nonlinear energy transfer in a gravity wave spectrum part 1: general theory". *Journal of Fluid Mechanism*, vol. 14, pp. 481-500, 1961.
- [HAS 85] HASSELMANN K., RANEY R., PLANT W., ALPERS W., SCHUMAN R., LYZENGA D., RUFENACH C., TUCKER M. "Theory of Synthetic Aperture Radar Ocean Imaging: A MARSEN view". *J. Geophys. Res.*, 90:4659–4686, 1985.
- [HAS 91] HASSELMANN K. HASSELMANN S. "On the nonlinear Mapping of an Ocean Wave Spectrum into a Synthetic Aperture Radar Image Spectrum and its Inversion". *J. Geophys. Res.*, 96:10713–10729, 1991.
- [HAS 96] HASSELMANN S., BRUNING C., HASSELMANNAND K., HEIMBACH P. "An improved algorithm for the retrieval of ocean wave spectra from synthetic aperture radar image spectra". *Journal of Geophysical Research*, vol 101, no. C7, pp. 16615-16629, 1996.
- [HEL 96a] HELLSTEN H., GUSTAVSSON A. "CARABAS, Airborne meter wave for ground surveillance". *Science et Défense 96, Paris*, 219–234. Dunod, 1996.
- [HEL 96b] HELLWICH O., MAYER H., WINKLER G. "Detection of Lines in Synthetic Aperture Radar (SAR) Scenes". *ISPRS, Int. Archives of Photogrammetry and Remote Sensing*, Vienna, 31:312–320, 1996.

- [HER 98] HERVET E., FJORTOFT R., MARTHON P., LOPÈS A. "Comparison of wavelet-based and statistical speckle filters". *SPIE EUROPTO'98*, Barcelona, Spain, 3497:43–54, 1998.
- [HER 00] HERVET E. "Appariement d'images radar à synthèse d'ouverture en conditions radargrammétiques". PhD Thesis, Institut National Polytechnique de Toulouse, 2000.
- [HIL 88] HILLION A., BOUCHER J.-M. "A new non linear filtering algorithm with application to radar images". *Proc. RADAR-88*, 177–181, 1988.
- [HIL 92] HILLION A., BOUCHER J.-M. "Filtrage des images radar (SAR)". *Statistique et analyse des données*, 16(2):35–57, 1992.
- [HOR 75] HORN B. *Obtaining Shape from Shading Information*. McGraw-Hill Book Company, 1975.
- [HOR 86] HORN B., BROOKS M. "The variational Approach to shape from shading". *Computer Vision, Graphics and Image Processing*, 33(2):174–208, 1986.
- [HOR 89] HORN B., BROOKS M. *Shape from Shading*. The MIT Press, Cambridge (USA), 1989.
- [HOR 93] HORAUD R., MONGA O. *Vision par ordinateur, Outils fondamentaux*. Hermès, Paris, 1993.
- [HUN 89] HUNTLEY J. "Noise-immune phase unwrapping algorithm". *Applied Optics*, 28(15):3268–3270, 1989.
- [ING 99a] INGLADA J., GARELLO R. "Depth estimation and 3D topography reconstruction from SAR images showing underwater bottom topography signatures". *IGARSS'99*, II, 956–958, Hamburg, Germany, 1999.
- [ING 99b] INGLADA J., GARELLO R. "Volterra model for the imaging mechanism of ocean surface currents by SAR". *IGARSS'99*, V:2251–2253, 1999.
- [ISH 78] ISHIMARU A. *Wave Propagation and Scattering in Random Media*. Academic Press, New York, 1978.
- [ISK 99] ISKANDER D., ZOUBIR A., BOASHASH B. "A method for estimating the parameters of the K distribution". *IEEE Trans. on IP*, 47(4):1147–1152, 1999.
- [JAC 75] JACKSON J. D. *Classical Electrodynamics*. John Wiley & Sons, 1975.
- [JAK 76] JAKEMAN E., PUSET N. "A model for non Rayleigh sea echo". *IEEE Trans. on AP*, 24(6):806–814, 1976.
- [JAK 84] JAKEMAN E. "Speckle statistics with a small number of scatterers". *Optical Engineering*, 23(4):453–461, 1984.
- [JAK 87] JAKEMAN E., TOUGH R. "Generalized K distribution: a statistical model for weak scattering". *J. Opt. Soc. Am.*, 4(9):1764–1772, 1987.
- [JAO 84] JAO J. "Amplitude distribution of composite terrain radar clutter and the K distribution". *IEEE Trans on AP*, AP-32(10):1049–1052, 1984.

- [JOU 93] JOUGHIN I., PERCIVAL D., WINNEBRENNER D. "Maximum Likelihood estimation of K-distribution parameters for SAR data". *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):989–999, 1993.
- [JUD 92] JUDGE T. R., QUAN C., BRYANSTON-CROSS P. J. "Holographic deformation measurements by Fourier transform technique with automatic phase unwrapping". *Optical Engineering*, 31(3):533–543, 1992.
- [KAY 81] KAY S. M., MARPLE S. L. "Spectrum analysis – a modern perspective". *IEEE Proceedings*, 69(11):1380–1419, 1981.
- [KEL 75] KELLER W., WRIGHT J. "Microwave Scattering and the Straining of Wind Generated Waves". *Radio Science*, 10:139–147, 1975.
- [KEL 88] KELLY P. A., DERIN H., HARTT K. D. "Adaptive Segmentation of Speckled Images Using a Hierarchical Random Field Model". *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-36(10):1628–1641, 1988.
- [KEY 82] KEYDEL W. "Application and experimental verification of an empirical backscattering cross section model of the Earth's surface". *IEEE Trans. on Geoscience and Remote Sensing*, GE-20(1):67–71, 1982.
- [KLE 92] KLEIN J. "Calibration of complex polarimetric SAR imagery using backscatter correlations". *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):183–194, 1992.
- [KON 87] KONG J., SWARTZ X., YUEH H., NOVAK L., SHIN R. "Identification of terrain cover using the optimum polarimetric classifier". *Journal of Electromagnetic Waves and Applications*, 2(2):171–194, 1987.
- [KRA 94] KRASITSKII V.P. "On the reduced equations in the theory of weakly nonlinear surface waves". *Journal of Fluid Mechanism*, vol. 272, pp. 1-20, 1994.
- [KRO 92] KROGSTAD H. "A simple Derivation of Hasselmann's Nonlinear Ocean-Synthetic Aperture Radar Transform". *J. Geophys. Res.*, 97:2421–2425, 1992.
- [KUA 85] KUAN D. T., SAWCHUK A. A., STRAND T. C., CHAVEL P. "Adaptive Noise Smoothing Filter for Images with Signal Dependant Noise". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7(2):165–177, 1985.
- [KUA 87] KUAN, SAWCHUK, STRAND, CHAVEL. "Adaptive Restoration of Images with Speckle". *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35(3):373–383, 1987.
- [LAC 88] LACROIX V. "A three modules strategy for edge detection". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:803–810, 1988.
- [LAP 63] LAPRADE G. "An analytical and experimental study of stereo for radar". *Photogram. Engineer.*, XXIX(2):294–300, 1963.
- [LAP 69] LAPRADE G. "Elevations from radar imagery". *Photogramm. Eng.*, XXXV:366–371, 1969.

- [LAU 89] LAUR H. "Analyse d'images radar en télédétection, discriminateurs radiométriques et texturaux". PhD Thesis, Université Paul Sabatier de Toulouse, n. 1403, 1989.
- [LAU 98] LAUR H., BALLY P., MEADOWS P., SANCHEZ J., SCHAEFTLER B., LOPINTO E., ESTEBAN D. Derivation of the backscattering coefficient σ^0 in ESA ERS SAR PRI products. European Space Agency (ES-TN-RS-PM-HL09), 1998.
- [LAV 97] LAVERGNAT J. SYLVAIN M. *Propagation des ondes radioélectriques*. Masson, 1997.
- [LE 97] LE CAILLEC J. *Etude et traitement des images SAR grâce aux moments et spectres d'ordres supérieurs*. ENST Bretagne – Thesis of the University of Rennes I, 1997.
- [LEB 86] LEBERL F. W., DOMIK G., RAGAM J., KOBRECK M. "Radar Stereomapping Techniques and Application to SIR-B Images of Mt. Shasta". *IEEE Transactions on Geoscience and Remote Sensing*, 24(4):473–481, 1986.
- [LEB 90] LEBERL F. *Radargrammetric Image Processing*. Artech House, Norwood, 1990.
- [LEC 89] LECHEVALIER F. *Principes de traitement des signaux radar et sonar*. Masson, Paris, 1989.
- [LEC 98] LECOZ D., DUPAS J., DUPLESSIS O. "Development status of the ONERA airborne SAR facilities (RAMSES)". *European Conference on Synthetic Aperture Radar*, 387–390. VDE, 1998.
- [LEC 02] LE CAILLEC J.M., GARELLO R. "Analysis of the SAR imaging process of the ocean surface using Volterra models". *IEEE Journal of Oceanic Engineering*, vol. 37, no. 3, pp. 675–699, July 2002.
- [LEE 81] LEE J.-S. "Refined Filtering of Image Noise Using Local Statistics". *Computer Graphics and Image Processing*, 15:380–389, 1981.
- [LEE 83] LEE J.-S. "Digital Image Smoothing and the Sigma Filter". *Computer Vision, Graphics and Image Processing*, 24:255–269, 1983.
- [LEE 91] LEE J.-S. "Speckle Reduction in Multipolarization, Multifrequency SAR Imagery". *IEEE Transactions on Geoscience and Remote Sensing*, TGARS-29(4):535–544, 1991.
- [LEE 94a] LEE J.-S., AINSWORTH T. L., GRUNES M. R., GOLDSTEIN R. M. "Noise filtering of interferometric SAR images". *Image and Signal Processing for Remote Sensing*, Rome, 2315, 735–742. SPIE-Europto, 1994.
- [LEE 94b] LEE J. S., HOPPEL K. W., MANGO S. A., MILLER A. R. "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery". *IEEE Trans. Geosci. and Remote Sensing*, 32(5):1017–1028, 1994.
- [LEE 94c] LEE J. S. JURKEVICH I. "Speckle filtering of synthetic aperture radar images: a review". *Remote Sensing Reviews*, 8:313–340, 1994.
- [LEH 96] LEHÉGARAT-MASCLE S. "Classification non supervisée d'images SAR polarimétriques". PhD Thesis, Télécom-Paris – ENST 96E024, 1996.

- [LET 92] LETOAN T., BEAUDOIN A., RIOM J., GUYON D. "Relating forest biomass to SAR data". *IEEE Transactions on Geoscience and Remote Sensing*, 30:403–411, 1992.
- [LI 90] LI F., GOLDSTEIN R. "Studies of multibaseline spaceborne interferometric synthetic aperture radars". *IEEE Trans. Geosci. and Remote Sensing*, 28(1):88–97, 1990.
- [LIN 92] LIN Q., VESECKY J. F., ZEBKER H. A. "New approaches in interferometric SAR data processing". *IEEE Trans. Geosci. and Remote Sensing*, 30(3):560–567, 1992.
- [LIO 91] LIONS P., ROUY E., TOURIN A. "Shape from Shading, Viscosity Solution and Edges". *Numerische Mathematik*, 1991.
- [LOF 97] LOFFELD O., ARNDT C. "Estimating the derivative of modulo-mapped phases". *ICASSP'97*, Munich, Germany, 2841–2844, 1997.
- [LOM 94] LOMBARDO P., OLIVER C. "Comparison of estimators for correlation properties of K-distributed SAR images". *SPIE E. C., Data Processing for Remote Sensing*, 2316, 1994.
- [LON 63] LONGUET-HIGGINS M. "The effect of non-linearities on statistical distributions in the theory of sea Waves". *J. of Fluids Mechanics*, 17:459–480, 1963.
- [LON 82] LONGUET-HIGGINS M. "On the skewness of sea-surface slopes". *J. of Physical Oceanography*, 12:1283–1291, 1982.
- [LON 83] LONG M. *Radar Reflectivity of Land and Sea*. Artech House, Inc, Washington, 1983.
- [LOP 90a] LOPÈS A., LAUR H., NEZRY E. "Statistical Distribution and texture in Multilook and Complex SAR Images". *IEEE Proc. of IGARSS'90*, 20–24, Washington, 1990.
- [LOP 90b] LOPÈS A., TOUZI R., NEZRY E. "Adaptive Speckle Filters and Scene Heterogeneity". *IEEE Transactions on Geoscience and Remote Sensing*, 28(6):992–1000, 1990.
- [LOP 93a] LOPÈS A. "De l'optique au Radar. Les applications de SPOT et ERS", 143–151. 1993.
- [LOP 93b] LOPÈS A., NEZRY E., TOUZI R., LAUR H. "Structure detection, and statistical adaptive filtering in SAR images". *Int. J. Remote Sensing*, 14(9):1735–1758, 1993.
- [LOP 97a] LOPÈS A., BRUNIQUEL J., SÉRY F., NEZRY E. "Optimal Bayesian texture estimators for speckle filtering of detected and polarimetric data". *Proc IGARSS'97*, Singapore, II:1044–1046, 1997.
- [LOP 97b] LOPÈS A., SÉRY F. "Optimal speckle reduction for the product model in multilook polarimetric data and the complex Wishart distribution". *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):632–647, 1997.
- [LYU 97] LYUBOSHENKO I., MAÎTRE H. "Robust Algorithms for Phase Unwrapping in SAR Interferometry". *SPIE Europto: Image Processing, and Synthetic Aperture Radar for Remote Sensing*, 3217, 176–187, London, UK, 1997.
- [LYU 98] LYUBOSHENKO I., MAÎTRE H. "Phase Unwrapping for Interferometric SAR Using Helmholtz Equation Eigenfunctions and the First Green's Identity". *Journal of the Optical Society of America A*, 16(2):378–395, 1998.

- [LYU 99] LYUBOSHENKO I., MARUANI A., MAÎTRE H. "Residue influence on phase unwrapping by use of Green-Helmholtz formulation". *Int. Geosc. and Remote Sensing Symposium: IGARSS'99*, III, 1537–1539, Hamburg, Germany, 1999.
- [LYU 00] LYUBOSHENKO I. MAÎTRE H. "Unwrapping circular interferograms". *Applied Optics*, 2000.
- [LYZ 02] LYZENGA D.R. "Unconstrained inversion of waveheight spectra from SAR images". *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 2, February 2002.
- [MAD 86] MADSEN S. N. "Speckle theory: modeling, analysis, and applications related to synthetic aperture data". PhD Thesis, Electromagnetic Institute, Technical University of Denmark, Lyngby, 1986.
- [MAD 87] MADSEN S. "Spectral properties of homogeneous and non homogeneous radar images". *IEEE Trans AES*, 23(4):583–588, 1987.
- [MAF 91] MAFFET A. WACKERMAN C. "The modified Beta density function as a model for SAR clutter statistics". *IEEE Transactions on Geoscience and Remote Sensing*, 29(2):277–283, 1991.
- [MAI 77] MAITRE H. "Propriétés fondamentales des signaux d'images". *Colloque National sur le Traitement du Signal et ses Applications*, Nice, 1977.
- [MAR 96] MARZOUKI A. "Segmentation statistique d'images radar". No. 1834, Lille University of Science and Technology, 1996.
- [MAR 97] MARINELLI L., TOUTIN T., DOWMAN I. "Génération de MNT par radargrammétrie: état de l'art et perspectives". *Bulletin de la Société Française de Photogrammétrie et Télédétection*, 148:88–96, 1997.
- [MAR 98] MARCOS S. *Les méthodes à haute résolution: traitement d'antenne et analyse spectrale*. Hermès, 1998.
- [MAS 93a] MASSONNET D., ROGRON B., CARMONA C. "Evolution des changements de phase de surface sur de grandes zones". *De l'optique au radar, les applications de SPOT et ERS*, 473–481. Cepadues Editions, 1993.
- [MAS 93b] MASSONNET D., ROSSI M., CARMONA C., PELTZER G., FEIGL K., ADRAGNA F. "The displacement field of the Landers earthquake mapped by radar interferometry". *Nature*, 364:138–142, 1993.
- [MAS 94] MASCLE S., OLIVIER P., VIDAL-MADJAR D. "Identification of vegetation classes using multi-configuration polarimetric SAR data". *Proceedings of Progress in Electromagnetic Research Symposium'94 (PIERS'94)*, 1994.
- [MAS 00] MASTENBROEK C., DE VALK C.F. "A semiparametric algorithm to retrieve ocean wave spectra from synthetic aperture radar". *Journal of Geophysical Research*, vol. 105, no. C2, pp. 3497–3516, 2000.
- [MCC 73] MCCORMICK M. *Ocean Engineering Wave Mechanics*. John Wiley & Sons Inc, New York, 1973.

- [MEN 63] MENON M. "Estimation of the shape and scale parameters of the Weibull distribution". *Technometrics*, 5:175–182, 1963.
- [MIL 74] MILLER K. *Complex Stochastic Process, an Introduction to Theory*. Addislon Wesley, Reading (MA; USA), 1974.
- [MOR 53] MORSE P. M., FESHBACH H. *Methods of Theoretical Physics*. McGraw-Hill, New York, 1953.
- [MUR 87] MURPHY H., HAYKIN S. "Bayesian classification of surface-based ice-radar images". *IEEE J. of Oceanic Engineering*, 12(3):493–503, 1987.
- [NAJ 97] NAJEH M., BEAULIEU J.-M. "A common evaluation approach to smoothing and feature preservation in SAR image filtering". *Proc. Intern. Symp. on Geomatic in the Era of Radarsat, GER'97*, Ottawa, Canada, 1997.
- [NEU 66] NEUMANN G., PIERSON W. *Principles of Physical Oceanography*. Prentice Hall, Englewood Cliffs, New Jersey, 1966.
- [NEZ 92] NEZRY E. "Restauration de la réflectivité radar, application à l'utilisation conjointe des images des radars à synthèse d'ouverture et optiques en télédétection". no. 1236, Paul Sabatier University, Toulouse, 1992.
- [NEZ 96] NEZRY E., LOPÈS A., D. DUCROT-GAMBART C. N., LEE J.-S. "Supervised Classification of K-distributed SAR images of natural targets and probability of error estimation". *IEEE Transactions on Geoscience and Remote Sensing*, 34(5):1233–1242, 1996.
- [NIC 97] NICOLAS J.-M., TUPIN F., SIGELLE M., MARUANI A. "Modélisation du chatoiement des images de Radar à Ouverture Synthétique à l'aide de la convolution de Mellin. Aspects théoriques". Ecole Nationale Supérieure des Télécommunications, ENST97D013, 1997.
- [NIC 98] NICOLAS J.-M., MARUANI A. "Speckle well Modeled by Mellin Transform: Application to SAR Images". *PIERS 1998*, 1, p. 383, Nantes, France, 1998.
- [NOC 96] NOCERA L. "Reconstruction du relief en imagerie radar à ouverture de synthèse: stéréoscopie et ombrages". PhD Thesis, University of Paris 7/Jussieu, 1996.
- [NOV 90] NOVAK L., BURL M. "Optimal speckle reduction in polarimetric SAR imagery". *IEEE Transactions on Aerospace and Electronic Systems*, 26(2):293–305, 1990.
- [OH 92] OH Y., SARABANDI K., ULABY F. "An empirical model and an inversion technique for radar scattering from bare soil surfaces". *IEEE Transactions on Geoscience and Remote Sensing*, 30:370–381, 1992.
- [OLI 86] OLIVER C. J. "The interpretation and simulation of clutter textures in coherent images". *Inverse Problems*, 2:481–518, 1986.
- [OLI 91] OLIVER C. J. "Information from SAR images". *J. Phys. D: Appl. Phys.*, 24:1493–1514, 1991.
- [OLI 93] OLIVER C. J. "Optimum texture estimators for SAR clutter". *J. Phys. D. Appl. Phys.*, 26:1824–1835, 1993.

- [OLI 96] OLIVER C. J., BLACKNELL D., WHITE R. G. "Optimum edge detection in SAR". *IEEE Proc. Radar Sonar and Navigation*, 143(1):31–40, 1996.
- [OLI 98] OLIVER C., QUEGAN S. *Understanding Synthetic Aperture Radar Images*. Artech House, Boston, London, 1998.
- [OLI 00] OLIVER C. "Rain Forest Classification Based on SAR Texture". *IEEE Trans on GRS*, 38(2):1095–1104, 2000.
- [OUC 87] OUCHI K., TAJBAKHSH S., BURG R. "Dependence of speckle statistics on back-scatter cross section fluctuations in SAR images of rough surfaces". *IEEE Transactions on Geoscience and Remote Sensing*, 25(5):623–628, 1987.
- [PAQ 97] PAQUERAULT S., MAÎTRE H. "Radarclinométrie opérationnelle". *Revue de la Société Française de Photogrammétrie et Télédétection*, 184:20–29, 1997.
- [PAQ 98] PAQUERAULT S. "Restitution du relief à partir d'images radar par radarclinométrie". PhD Thesis, École Nationale Supérieure des Télécommunications, ENST 98E007, 1998.
- [PEN 94] PENICAND C., CHAMPSEIX N., LEDEZ C., LENTZ C., RUDANT J.-P. "Utilisation de l'imagerie satellitaire radar pour la cartographie de base en milieu tropical". *5th European Conf. and Exhib. on Geogr. Inf. Systems (EGIS)*, Paris, 1994.
- [PER 95] PERLANT F. "Utilisation de l'interférométrie radar pour la réalisation et l'interprétation d'ortho-images". *Journées Cartographie Radar*, Nice, 138, 54–65. Bulletin SFPT, 1995.
- [PHI 77] PHILLIPS O. *Dynamics of the Upper Ocean* (2nd edition). Cambridge University Press, Cambridge, 1977.
- [PIE 94] PIERCE L., ULABY F., SARABANDI K., DOBSON M. "Knowledge-based classification of polarimetric SAR images". *IEEE Transactions on Geoscience and Remote Sensing*, 32(5):1081–1086, 1994.
- [PLA 86] PLANT W. "A two-scale Model of Short Wind-Generated Waves and Scatterometry". *J. Geophys. Res.*, 91:10 735–10749, 1986.
- [POL 97] POLIDORI L. *Cartographie Radar*. Gordon and Breach Science Publishers, Amsterdam, Holland, 1997.
- [POS 93] POSNER F. "Texture and speckle in high resolution SAR clutter". *IEEE Transactions on Geoscience and Remote Sensing*, 31(1):192–203, 1993.
- [PRA 75] PRATT W. K. "Vector space formulation of two dimensional signal processing operations". *Computer Graphics and Image Processing*, 4:1–24, 1975.
- [PRA 90a] PRATI C., GIANI M., LEURATTI N. "SAR Interferometry: a 2D phase unwrapping technique based on phase and absolute values informations". *IGARSS'90*, Washington DC, USA, 3, 2043–2046, 1990.
- [PRA 90b] PRATI C., ROCCA F., GUARNIERI A., DAMONTI E. "Seismic migration for SAR focusing: interferometrical application". *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):627–639, 1990.

- [PRA 94] PRATI C., ROCCA F., GUARNIERI A. M., PASQUALI P. “Reports on ERS-1 SAR interferometric techniques and applications”. Politecnico di Milano, 1994.
- [PRI 94] PRITT M. D., SHIPMAN J. S. “Least-squares two-dimensional phase unwrapping using FFTs”. *IEEE Transactions on Geoscience and Remote Sensing*, 32(3):706–708, 1994.
- [PRI 95a] PRITT M. D. “Unweighted least squares phase unwrapping by means of multigrid techniques”. *Synthetic Aperture Radar and Passive Microwave Sensing*, Paris, 2584 *EUROPTO*, 289–300. SPIE, 1995.
- [PRI 95b] PRITT M. D. “Weighted least squares phase unwrapping by means of multigrid techniques”. *Synthetic Aperture Radar and Passive Microwave Sensing*, Paris, 2584 *EUROPTO*, 278–288. SPIE, 1995.
- [QUE 90] QUELLE H.-C., BOUCHER J.-M., PIECZYNSKI W. “Unsupervised Bayesian Classification of SAR Images”. *ISPRS*, 1990.
- [QUE 93a] QUELLE H.-C. “Segmentation bayésienne non supervisée en imagerie radar”. PhD Thesis, University of Rennes I, no. 950, 1993.
- [QUE 93b] QUELLE H.-C., DELIGNON Y., PIECZYNSKI W. “Unsupervised Bayesian segmentation of SAR images using the Pearson system distributions”. *IGARSS'93*, Tokyo, 18-21:1538–1540, 1993.
- [QUE 94] QUEGAN S. “A unified algorithm for phase and crosstalk calibration – theory and observations”. *IEEE Transactions on Geoscience and Remote Sensing*, 32(1):89–99, 1994.
- [QUE 95] QUEGAN S., RHODES I. “Statistical models for polarimetric data: consequences, testing and validity”. *International Journal of Remote Sensing*, 16:1183–1210, 1995.
- [QUE 97] QUELLE H.-C., BOUCHER J.-M., PIECZYNSKI W. “Adaptive parameter estimation and unsupervised image segmentation”. *Machine Graphics and Vision*, 5(4):613–631, 1997.
- [RAM 86] RAMAPRIYAN H., STRONG J., HUNG Y., MURRAY C. “Automated matching of pairs of SIR-B images for elevation mapping”. *IEEE Transactions on Geoscience and Remote Sensing*, TGARS-4(4):462–472, 1986.
- [RAN 80] RANEY R. “SAR response to partially coherent phenomena”. *IEEE trans. On Antennas and Propagation*, 28:777–787, 1980.
- [RAN 88] RANEY R., VESSEL G. “Spatial consideration in SAR speckle simulation”. *IEEE Transactions on Geoscience and Remote Sensing*, 26(5):666–671, 1988.
- [RAN 93] RANGASWAMY M., WEINER D., OZTURK A. “Non-gaussian random vector identification using spherically invariant random processes”. *IEEE Trans. on Aerospace and Electron. Systems*, 29(1):111–124, 1993.
- [REI 86] REID G. T. “Automatic Fringe pattern analysis: a review”. *Optics and Lasers in Engineering*, (7):37–68, 1986.
- [RIG 91] RIGNOT E., CHELLAPPA R., KWOK R. “Classification of Multifrequency Multilook Synthetic Aperture Radar Data”. *IEEE Conference*, 1801–1804, 1991.

- [RIG 92] RIGNOT E., CHELLAPPA R. "Segmentation of polarimetric synthetic aperture radar data". *IEEE Transactions on Image Processing*, 1(3):281–299, 1992.
- [RIG 93] RIGNOT E., KWOK R. "Characterisation of spatial statistics of distributed targets in SAR data". *Int J. of Remote Sensing*, 14(2):345–363, 1993.
- [ROD 92] RODRIGUEZ E., MARTIN J. "Theory and design of interferometric synthetic aperture radars". *IEEE PROCEEDINGS-F*, 139(2):147–159, 1992.
- [ROD 97] RODENAS J., GARELLO R. "Wavelet Analysis in SAR Ocean Image Profiles for Internal Waves Detection and Wavelength Estimation". *IEEE Trans. on Geoscience and Remote Sensing*, 35(4):933–945, 1997.
- [ROD 98] RODENAS J., GARELLO R. "Internal wave detection and location in SAR images using wavelet transform". *IEEE Trans. on Geoscience and Remote Sensing*, 36(5):1494–1507, 1998.
- [ROD 99] RODENAS J. *Détection et localisation d'ondes internes sur des images SAR de la surface de l'océan*. ENST Bretagne – Thesis from the University of Rennes I, 1999.
- [ROM 97] ROMEISER R., ALPERSW. "An Improved Composite Surface Model for the Radar Backscattering Cross Section of the Ocean Surface. 2. Model Response to Surface Roughness Variations and the Radar Imaging of Underwater Bottom Topography". *Journal of Geophysical Research*, 102(25):251–267, 1997.
- [ROS 68] ROSENFIELD G. H. "Stereo Radar Techniques". *Photogram. Engineer.*, XXXIX:586–594, 1968.
- [RUD 96] RUDANT J., BALTZER F., DEROIN J., LOINTIER M., MAÎTRE H., MOUGIN E., PÉNICAND C., PROST M. "Apport des images radar satellitaires ERS11 et JERS1 dans le domaine de la cartographie générale et thématique en contexte tropical humide". *SFPT Société Française de Photogrammétrie et Télédétection*, (142):15–31 and 81–82, 1996.
- [SAF 89] SAFA F., FLOUZAT G. "Speckle removal on radar imagery based on mathematical morphology". *Signal Processing*, 16:319–333, 1989.
- [SCH 94] SCHMITT M., MATTIOLI J. *Morphologie Mathématique*. Masson, Paris, France, 1994.
- [SCH 96] SCHULER D., LEE J., GRANDI G. D. "Measurement of topography using polarimetric SAR images". *IEEE Trans on Geoscience and Remote Sensing*, 34(5):1266–1277, 1996.
- [SCH 98] SCHULER D., AINSWORTH T., LEE J., GRANDI G. D. "Topography mapping using Polarimetric SAR data". *International Journal of Remote Sensing*, 19(1):141–160, 1998.
- [SCH 00] SCHULER D., LEE J., AINSWORTH T. "Measurement/compensation of terrain slopes using multi-pass polarimetric SAR data". VERLAG V., *3rd European Conference on Synthetic Aperture Radar*, 1, 269–272, 2000.
- [SER 97] SERY F. "Filtrage du speckle dans les images RSO polarimétriques et classification non supervisée multi-sources". PhD Thesis, University of Paul Sabatier (Toulouse III) no. 1236, 1997.

- [SHA 94] SHAHSHANI B., LANDGREBE D. "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon". *IEEE Transactions on Geoscience and Remote Sensing*, 32(5):1087–1095, 1994.
- [SHE 92] SHEN J., CASTAN S. "An optimal linear operator for step edge detection". *CVGIP*, 54(2):112–133, 1992.
- [SHI 96] SHISTAD-SOLBERG A., TAXT T., JAIN A. "A Markov random fields model for classification of multisource satellite imagery". *IEEE Trans on Geoscience and Remote Sensing*, 34(1):100–113, 1996.
- [SIJ 98] SIJBERS J., DEN DEKKER A., SCHEUNDERS P., VAN DYCK D. "Maximum-Likelihood estimation of Rician distribution parameters". *IEEE Trans. on Med. Imaging*, 17(3):357–361, 1998.
- [SMI 97] SMITS P. C., DELLEPIANE S. G. "Synthetic Aperture Radar Image Segmentation Detail Preserving Markov Random Field Approach". *IEEE Transactions on Geoscience and Remote Sensing*, 35(4):844–857, 1997.
- [SPA 93] SPAGNOLINI U. "2-D phase unwrapping and phase aliasing". *Geophysics*, 58(9):1324–1334, 1993.
- [SPA 95] SPAGNOLINI U. "2-D Phase unwrapping and instantaneous frequency estimation". *IEEE Trans. Geosci. and Remote Sensing*, 33(3):579–589, 1995.
- [SPI 91] SPIK A., ROBINSON D.W. "Investigation of the cellular automata method for phase unwrapping and its implementation on an array processor". *Optics and Lasers in Engineering*, 14:25–37, 1991.
- [STE 67] STERN J., BARBEYRAC J. D., POGGI R. *Méthodes pratiques d'étude des fonctions aléatoires*. Dunod, 1967.
- [STO 89] STOICA P., NEHORAI A. "MUSIC, Maximum Likelihood, and Cramer-Rao Bound". *IEEE Trans. Acoustic, Speech and Signal Processing*, 37(5):720–741, 89.
- [SUC 00] SUCHAIL J., BUCK C., GUIJARRO J., SCHÖNENBERG A., TORRES R. "The Envisat ASAR instrument". VERLAG V., *3rd European Conference on Synthetic Aperture Radar*, 1, 33–36, 2000.
- [SYL 97] SYLVANDER S., COUSSON D., GIGORD P. "Etude des performances géométriques des images RadarSat". *Bull. Société Française de Photogrammétrie et Télédétection SFPT*, 148:57–65, 1997.
- [SYL 98] SYLVANDER S., PETITCOLAS B., COUSSON D., SOUYRIS J.-C. "Evaluation of the potential of RadarSat images for the production of Digital Terrain Model by radargrammetry". *Final ADRO Symposium*, Montreal Canada, 1998.
- [SZA 77] SZAJNOWSKI W. "Estimators of log-normal distribution parameters". *IEEE Trans. on Aerospace and Electr. Systems*, 13(5):533–536, 1977.
- [TAK 88a] TAKAJO H., TAKAHASHI T. "Least-squares phase estimation from the phase difference". *Journal of the Optical Society of America A*, 5:416–425, 1988.

- [TAK 88b] TAKAJO H., TAKAHASHI T. "Noniterative method for obtaining the exact solution for the normal equation in least-squares phase estimation from the phase difference". *J. Opt. Soc. Am.*, 5(11):1818–1827, 1988.
- [TAR 96] TARAYRE H. "Extraction de modèles numériques de terrain par interférométrie radar satellitaire: algorithmie et artefacts atmosphériques". PhD Thesis, Institut National Polytechnique de Toulouse, 1996.
- [THO 91] THOMAS J., KOBER W., LEBERL F. "Multiple-image SAR shape from shading". *Photogrammetric Engineering and Remote Sensing*, 57(1):51–59, 1991.
- [TOU 88a] TOUZI R. "Analyse d'images radar en télédétection: améliorations radiométriques, filtrage du speckle et détection des contours". PhD Thesis, University of Paul Sabatier, Toulouse, no. 258, 1988.
- [TOU 88b] TOUZI R., LOPÈS A., BOUSQUET P. "A Statistical and Geometrical Edge Detector for SAR Images". *IEEE Transactions on Geoscience and Remote Sensing*, 26(6):764–773, 1988.
- [TOU 92] TOUZI R., GOZE S., LETOAN T., LOPES A., MOUGIN E. "Polarimetric discriminators for SAR images". *IEEE Transactions on Geoscience and Remote Sensing*, 30(5):973–980, 1992.
- [TOU 95] TOUGH R., BLACKNELL D., QUEGAN S. "A statistical description of polarimetric and interferometric synthetic aperture radar data". *Proc. Roy. Soc. London A*, 449:567–569, 1995.
- [TOU 96] TOUZI R., LOPÈS A. "Statistics of the Stokes parameters and the complex coherence in 1-look and multi-look speckle fields". *IEEE Transactions on Geoscience and Remote Sensing*, 34(2):519–531, 1996.
- [TOU 99] TOUZI R., LOPÈS A., BRUNIQUEL J., VACHON P. "Coherence estimation of SAR imagery". *IEEE Transactions on Geoscience and Remote Sensing*, 37(1):135–149, 1999.
- [TOU 00] TOUTIN T., GRAY L. "State-of-the-art of elevation extraction from satellite SAR data". *Photogrammetry and Remote Sensing*, 55(1):13–33, 2000.
- [TOW 89] TOWERS D. P., JUDGE T. R., BRYANSTON-CROSS P. J. "A quasi heterodyne holographic technique and automatic algorithms for phase unwrapping". *Fringe Pattern Analysis*, 1163, 95–119. SPIE, 1989.
- [TRO 96a] TROUVÉ E. "Imagerie interférentielle en radar à ouverture synthétique". PhD Thesis, Télécom-Paris ENST 96E014, 1996.
- [TRO 96b] TROUVÉ E., CARAMMA M., MAÎTRE H. "Fringe detection in noisy complex interferograms". *Applied Optics*, 35(20):3799–3806, 1996.
- [TRO 98] TROUVÉ E., NICOLAS J.-M., MAÎTRE H. "Improving Phase Unwrapping Techniques by the Use of Local Frequency Estimates". *IEEE Trans. on Geoscience and Remote Sensing*, 36(6):1963–1972, 1998.
- [TSA 85] TSANG L., KONG J. A., SHIN R. *Theory of Microwave Remote Sensing*. Wiley Interscience, New York, 1985.

- [TUC 83] TUCKER M. "The decorrelation Time of microwave echoes from the sea surface". *International Journal of Remote Sensing*, 6:1075–1089, 1983.
- [TUP 96] TUPIN F., TROUVÉ E., DESCOMBES X., NICOLAS J.-M., MAÎTRE H. "Improving IFSAR Phase Unwrapping by Early Detection of Non-interferometric Features". *SPIE/EUROPTO European Symposium on Satellite Remote Sensing III*, 2958, 250–261, Taormina, Italy, 1996.
- [TUP 97] TUPIN F., SIGELLE M., CHKEIF A., VERAN J.-P. "Restoration of SAR images using recovery of discontinuities and non-linear optimization". *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, Venice, Italy, 1997.
- [TUP 98] TUPIN F., MAÎTRE H., MANGIN J.-F., NICOLAS J.-M., PECHERSKY E. "Linear Feature Detection on SAR Images: Application to the Road Network". *IEEE Transactions on Geoscience and Remote Sensing*, 36(2):434–453, 1998.
- [TUP 99] TUPIN F., BLOCH I., MAÎTRE H. "A First Step towards Automatic Interpretation of SAR Images using Evidential Fusion of Several Structure Detectors". *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1327–1343, 1999.
- [TUP 00] TUPIN F. "Radar cross-views for road detection in dense urban areas". *EUSAR 2000*, Munich, 2000.
- [ULA 81] ULABY F., MOORE R., FUNG A. *Microwave Remote Sensing*. Addison Wesley, Reading (USA), 1981.
- [ULA 82] ULABY F. T., MOORE R. K., FUNG A. K. *Microwave Remote Sensing*, volume II. Artech House, 1982.
- [ULA 86a] ULABY F., DOBSON M. *Handbook of Radar Scattering Statistics for Terrain*. Artech House Inc, Boston (Mi; USA), 1986.
- [ULA 86b] ULABY F. T., KOUYATE F., BRISCO B., WILLIAMS T. L. "Textural Information in SAR Images". *IEEE Transactions on Geoscience and Remote Sensing*, 24(2):235–245, 1986.
- [ULA 87] ULABY F., HELD D., DOBSON M., MCDONALD K., SENIOR T. "Relating polarization phase difference of SAR signals to scene properties". *IEEE Transactions on Geoscience and Remote Sensing*, 25(1):83–91, 1987.
- [ULA 90] ULABY F. ELACHI C. *Radar Polarimetry for Geoscience Applications*. Artech House, Boston, London, 1990.
- [VAN 77] VANDERBRUG G. J. "Experiments in iterative enhancement of linear features". *Computer Graphics and Image Processing*, 6:25–42, 1977.
- [VAN 87] VAN ZYL J., ZEBKER J., ELACHY H. "Imaging radar polarisation signatures: theory and observation". *Radio Science*, 22(4):529–543, 1987.
- [VAN 89] VAN ZYL J. "Unsupervised classification of scattering behavior using radar polarimetry data". *IEEE Transactions on Geoscience and Remote Sensing*, 27(1):36–45, 1989.

- [VAN 90] VAN ZYL J. "Calibration of polarimetric radar images using only images parameters and trihedral corner reflector responses". *IEEE Transactions on Geoscience and Remote Sensing*, 28:337–348, 1990.
- [VAN 91] VAN ZYL J., BURNETTE C. "Data volume reduction for single look polarimetric imaging radar data". *IEEE Transactions on Geoscience and Remote Sensing*, 29(5), 1991.
- [WAL 99] WALESSA M. "Texture preserving despeckling of SAR images using GMRFs". *IGARSS'99*, 1999.
- [WER 00] WERNER M. "Shuttle Radar Topographic Mission (SRTM) Mission Overview". VERLAG V., *3rd European Conference on Synthetic Aperture Radar*, 1, 209–212, 2000.
- [WHI 91] WHITE R. G. "Change detection in SAR imagery". *International Journal of Remote Sensing*, 12(12):339–360, 1991.
- [WHI 94] WHITE R. "A simulated annealing algorithm for SAR andMTI image cross-section estimation". *SPIE SAR Data Processing for Remote Sensing*, 2316:137–145, 1994.
- [WIL 86a] WILDEY R. "Radarclinometry". *Earth, Moon and Planets*, 36:217–247, 1986.
- [WIL 86b] WILDEY R. "Radarclinometry for the Venus Radar Mapper". *Photogrammetric Engineering and Remote Sensing*, 51(1):41–50, 1986.
- [WRI 66] WRIGHT J. "Backscattering from Cappillary Waves with Application to sea Clutter". *IEEE trans. on Antennas and Propagation*, AP-14:749–754, 1966.
- [WRI 68] WRIGHT J. "A New Model for Sea Clutter". *IEEE trans. on Antennas and Propagation*, AP-16:217–223, 1968.
- [WU 92] WU Y., MAÎTRE H. "Smoothing speckled synthetic aperture radar images by using maximum homogeneous region filters". *Optical Engineering*, 31(8):1785–1792, 1992.
- [XIE 99] XIE H., ULABY F. T., PIERCE L. E., DOBSON M. C. "Performance Matrix for SAR speckle suppression filters". *IGARSS'99*, Hamburg, Germany, 1999.
- [YUE 89] YUEH S., KONG J., JAO J., SHIN R., NOVAK L. "K-distribution and polarimetric radar clutter". *Journal of Electromagnetic Waves and Applications*, 8:747–768, 1989.
- [YUE 90] YUE S., KONG J., JONG J., SHIN R., ZEBKER H., LETOAN T., OTTL H. K Distribution and polarimetric radar clutter. KONG J., *Polarimetric Remote Sensing*. Elsevier, Amsterdam, 1990.
- [ZAK 68] ZAKHAROV V.E. "Stability of periodic waves of finite amplitude on the surface of a deep fluid". *Appl. Mech. Tech. Phys.* (English translation), 2, pp. 190-194, 1968.
- [ZEB 98] ZEBKER H. A. LU Y. "Phase unwrapping for radar interferometry: residue-cut, least-squares, and synthesis algorithms". *J. Opt. Soc. Am. A*, 15:586–598, 1998.
- [ZRN 75] ZRNIC D. S. "Moments of estimated input power for sample averages of radar receiver outputs". *IEEE Transactions on Aerospace Electronics Systems*, 11(1):109–113, 1975.

List of Authors

Frédéric ADRAGNA

CNES

Toulouse

Danielle DUCROT

CESBIO-CNES

Toulouse

René GARELLO

IT-TELECOM Bretagne

Brest

Jean-Marc LE CAILLEC

IT-TELECOM Bretagne

Brest

Sylvie LE HÉGARAT-MASCLE

University of Paris XI

Orsay

Armand LOPÈS

CESBIO-CNES

Toulouse

Henri MAÎTRE

TSi Department

TELECOM ParisTech

Paris

Jean-Marie NICOLAS
TSi Department
TELECOM ParisTech
Paris

Emmanuel TROUVÉ
LISTIC
Polytech'Savoie
Annecy

Florence TUPIN
TSi Department
TELECOM ParisTech
Paris

Index

A

ACF, 307
ACM, 308, 310
Aer-1, 201
albedo, 13, 255
algorithm
 EM, 180, 184
 SEM, 181, 186
aliasing, 46, 304, 310, 313
along track, 32, 51, 71, 85
Lambertian reflectance, 262, 264
ambiguity height, 288, 289, 291, 293
apodization, 27, 48
attitude, 233
azimuth, 27, 41–54, 76–79, 85, 187, 194, 196, 235
azimuth smearing, 332, 333, 345, 347, 349, 352, 353

B

B/H, 288, 289
backscattering, 10, 87, 193, 272
band
 Doppler, 45, 47
 frequency, 44, 49, 50
bathymetry, 331, 333, 352, 353
Bean’s atlas, 9, 298
bias, 85, 104, 112, 118, 120–124, 147, 148, 153, 169–171, 239, 271, 299, 324
bi-static, 281, 290

Bragg wavelength, 337
BSA convention, 18, 19

C

calibration, 75, 83–85, 93, 139, 141, 265
CARABAS, 73
Cauchy’s theorem, 320
CEOS format, 76
CFAR, 208, 217
channel imbalance, 85
chirp, 35, 38, 40, 42, 46, 51, 83, 102
classification
 Bayesian, 198, 201
 contextual, 181, 189–191, 317
 entropy-based, 195
 point, 181, 189, 192
clutter, 125, 126, 130, 132–135, 222, 268, 300
coefficient of variation, 95, 96, 107–111, 128, 131, 148, 159–165, 187–192, 213, 215
coherent illumination, 14
co-index, 9, 298
complex multilooking, 313
contrast, 87, 95, 97, 122–124, 207, 212, 215–217, 226, 228, 240, 241, 317, 332, 333
copolar, 82, 85, 139
Crimmins, 144
critical baseline, 288, 295, 296
cross-polar, 82, 85, 140, 141, 163, 194, 197, 198
cross-talk, 85

cross track, 32, 51, 71, 85
cut, 140, 320, 321, 329

D

dating, 235, 239
degree of coherence, 98, 111, 116, 118, 285
derived channels, 187, 189, 109, 192
detection
 contour 220, 223, 226
 edge, 182
 line, 229
 target, 125, 130, 169, 208, 210, 213,
 214, 218, 225
DIAPASON, 313
differential interferogram, 292, 325, 327
digamma, 131
disparity, 248–257
Doppler
 centroid, 236
 frequency, 44, 49, 50
 zero, 53, 236, 240

E

electronic density, 8, 299
entropy, 195, 197, 198
epipolar line, 245, 251
equation
 eikonal, 7, 9, 262, 265, 278
 Fredholm, 328
 Poisson, 323–325
 radar, 33, 34, 83, 90
ergodicity, 308, 345
estimation-maximization, 180

F

fading, 94
far range, 31, 48, 51, 54, 77, 78, 235
filter
 Frost, 152, 153, 170, 172, 253
 Kuan, 151, 152, 160, 169, 172, 189, 192
 Lee, 151, 152, 313
 Wiener, 153
filtering
 multi-temporal, 187
 polarimetric, 161

first Green identity, 319
Fisher-Tippett, 97, 112
foreshortening, 54
FRINGE, 312
fringes,
 orbital, 287
 topographic, 287, 293
function
 autocorrelation, 99, 100, 307, 346
 Helmholtz eigenfunction, 328
 modulation transfer, 337, 340, 347
 transfer, 45, 335, 340, 341, 343, 353
fusion, 191, 192, 228, 241, 317, 319

G

Gamma
 generalized, 97, 104, 107–109, 134, 187
 incomplete, 108
gate, 52–55, 93, 101–105, 111, 115
geometry
 epipolar, 245, 246, 249, 289
 radar, 40, 53, 54, 239, 271
ghost line, 320, 321
gradient, 214, 223, 225, 267, 269, 272, 278,
 305, 306, 312, 323–325, 329
Green, 319, 326, 328
ground swath, 25, 31
ground-truth, 176, 179, 187

H, I

Huynen decomposition, 195, 196
hydrometeors, 10
IFSARE, 291
internal wave, 331, 332
image
 coherence, 281, 282, 315, 317
 mirror, 329
 scope, 29
 single-look, 87, 107, 110, 135, 169, 171,
 211
inversion, 267, 268, 271, 278, 333, 345,
 349, 352–355
ionosphere, 8, 9, 35, 73, 235, 297–299
iso-elevation, 289

J–L

Jones representation, 23
 Kuan, 151, 160, 168, 169, 172, 183, 189, 192
 Lambert, 262, 264, 269, 272,
 Laplace, 96, 97
 law
 complex Wishart, 137
 Descartes, 11, 14
 Fisher-Tippett, 97, 112
 Gamma, 111, 150, 155, 168, 179, 201,
 210, 211, 218, 253
 Goodman, 154, 218
 Lambert, 272
 Laplace, 96, 97
 lognormal, 134
 mixed, 184
 Nagakami-Rice, 134, 135
 Snell-Descartes, 11, 12
 Weibull, 134
 least squares, 155, 239, 252, 304, 310, 312,
 319, 322–325
 Lee, 144, 151, 152, 168, 172, 313
 Lin, 287
 LMMSE, 146, 160

M

MARSEN, 345
 mask, 225, 228, 304, 321, 322, 325
 matching, 244–258, 337
 matrix
 autocorrelation, 306, 308
 backscattering, 19, 20, 22, 196
 coherence, 195, 196
 covariance, 20, 98, 104, 110, 115, 127,
 136–138, 140, 143, 148, 160–171,
 182–184, 195–201, 219, 220, 283
 Mueller, 19–23, 163, 196
 Stokes, 21, 81, 82
 maximum
 a posteriori (MAP), 146, 147, 156, 178
 likelihood (ML), 96, 146, 165, 178, 187,
 191, 208, 215, 307
 mesosphere, 9
 minimum mean square error, 147
 MMIC, 27

MMSE, 146, 147, 151, 154, 159, 166, 171,
 183, 187–192, 313
 model
 digital terrain model (DTM), 353
 product, 125, 126, 137, 138, 141, 161,
 165, 201
 MTF, 337, 343
 Mueller, 19–23, 82, 163, 165, 196
 multi-look, 75, 78, 80, 89, 94–99, 105–120,
 135, 137, 144, 148–150, 163–165, 183–
 189, 200, 213, 217, 283, 306
 multi-scale, 225, 231, 306, 310–312
 MUSIC, 306, 309, 310

N

Nadir, 47–49, 70
 NASA, 62, 76, 80
 NASDA, 64, 71, 76
 near range, 31, 48, 51, 54, 77, 78
 Neumann condition, 323
 non-linearities, 339, 348, 352, 354

O

orbit
 ascending, 241, 247, 299
 descending, 60, 240–242, 247, 299
 restored, 234
 overlay, 26, 54

P

Processing and Archiving Facility (PAF),
 78, 83
 permeability, 2, 3
 permittivity, 2, 3, 6, 7, 12
 phase
 main, 302
 specific, 63, 281
 unwrapped, 302, 323, 326
 wrapped, 282, 301, 305, 306, 312, 324,
 326, 329
 phase unwrapping, 301
 photogrammetry, 243, 245, 251, 257
 Poincaré sphere, 23
 polarization
 circular, 16, 17, 194

elliptical, 16
 linear, 16, 17
 PRF, 30, 32, 46–53, 63–68, 77, 236, 237
 PRI, 78, 79, 83
 propagation speed, 335
 proximal distance, 234, 235, 239, 240
 pulse compression, 35, 38, 40, 47, 66, 75, 76, 102
 pulse repetition frequency, 30, 32, 46, 64, 77
 pyramid, 254

R

radar cross-section (RCS), 13, 33, 34, 36, 83, 93, 135, 171
 radar reflectivity, 83, 87, 89, 92–99, 104, 114, 124, 126, 129
 radar shadow, 55
 radarclinometry, 261
 radarpolarimetry, 276, 278
 radiometry, 114, 126, 152, 170, 176, 208, 212, 215, 217, 225, 227, 241, 255, 256, 262, 265, 266, 271, 280, 321
 random walk, 89–91, 135
 RAR, 25, 50, 337, 344–349, 354
 Rayleigh, 10–13, 92, 95, 97, 104, 109, 135
 reciprocity principle, 19, 21
 registration, 233, 237, 239, 240, 242, 281, 282
 regularization, 152, 186, 228, 244, 266–269, 319, 352–355
 relation
 Lin, 287
 Siegert, 100, 141, 164
 residue, 304, 313, 315, 319–324, 329
 ripple, 337
 ROC, 217
 Rayleigh criterion, 13, 92

S

SAR, 25
 scanning, 239, 337
 scanning distortion, 337
 segmentation, 88, 105, 125, 128, 134, 143, 144, 169, 170, 171, 181–192, 228, 240, 319
 SEM, 181, 184, 186

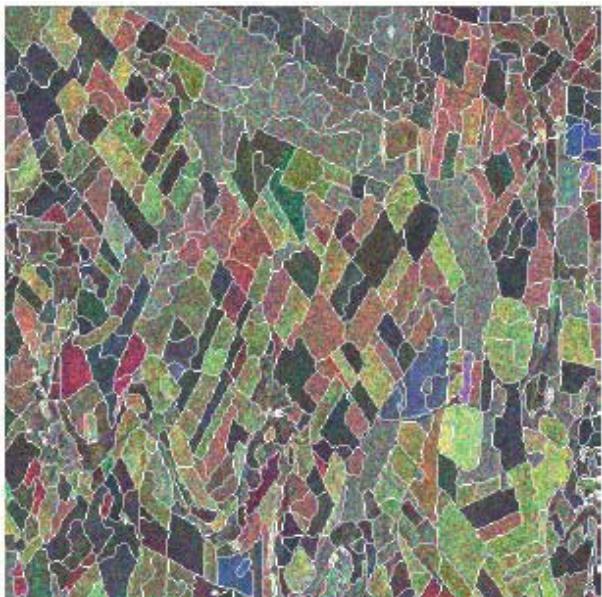
shape from shading, 261–263
 Siegert, 100, 141, 164
 sigma nought, 83
 signal-to-noise ratio (SNR), 34, 104, 299, 300
 single-look complex (SLC), 75–83, 97, 98, 101, 104, 105, 114, 115, 154, 187, 231, 313
 slant range, 53
 Smith-Weintraub equation, 9
 Smith-Weintraub formula, 298
 spatial triangulation, 242
 spectral estimation, 315
 spectrum splitting, 348
 SRTM, 66, 68, 69, 279, 285, 290, 291
 stationarity, 310, 345
 stereoscopic base, 244, 246, 248
 stereoscopy, 243
 stereovision, 243, 244, 246, 254, 258, 288, 289
 Stokes, 17, 21–23, 81, 82, 89, 324, 334
 vector, 17, 21, 23
 stratosphere, 9
 strong reflector, 207, 209

T

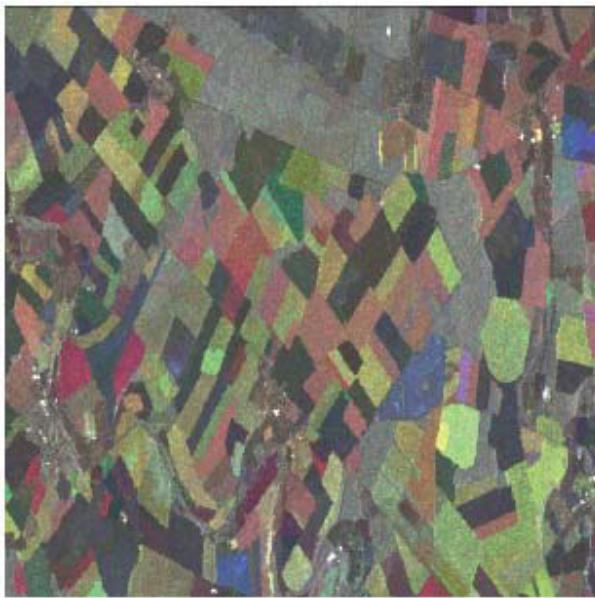
tandem coherence, 293
 temporal series, 171, 176, 183, 186, 191, 241
 texture, 95, 125–129, 133, 137, 138, 141, 144, 148, 153–172, 176, 180, 187, 190, 191, 201, 208, 256, 258
 tilt modulation, 337, 340–342
 traveling wave tube (TWT), 35
 troposphere, 9, 35, 235, 297–299

V–Z

vectorial, 90, 98, 115, 160, 161, 170, 171, 183, 187–192, 196, 303, 313, 323
 velocity bunching, 332, 333, 344, 347–349
 Volterra model, 345, 348–354
 weighted least squares, 325, 327
 Wishart, 116, 137, 165, 182, 183, 187, 191, 200–203
 zero-padding, 103



(a) Original image (colored composition) and segmentation



(b) Filtered image

Figure 7.2. Colored composition of images from 3 dates obtained by means of the linear vector MMSE estimator (b) where the statistical parameters are calculated over entire regions produced by radiometric segmentation (a)

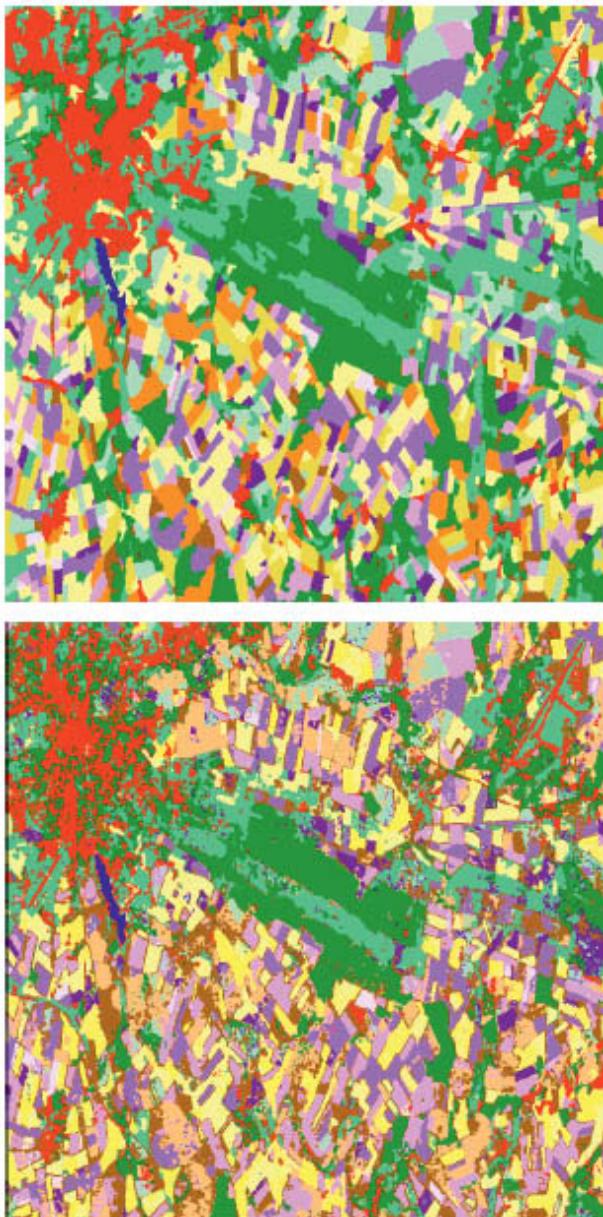
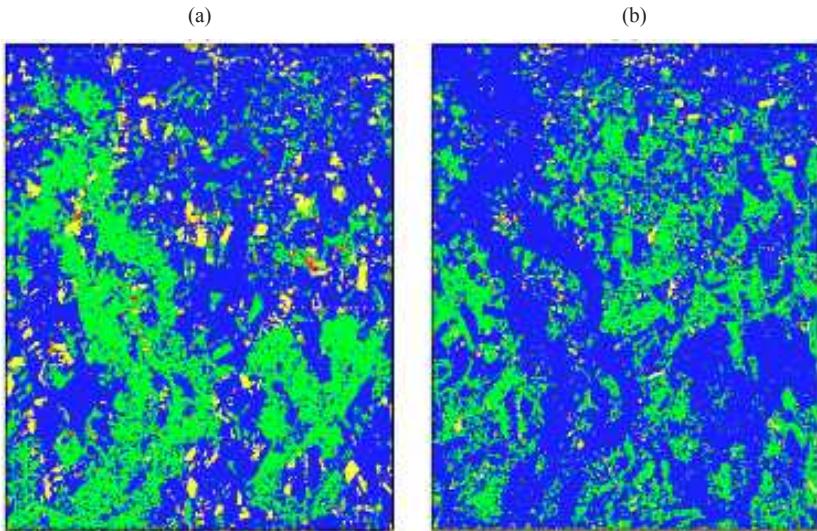
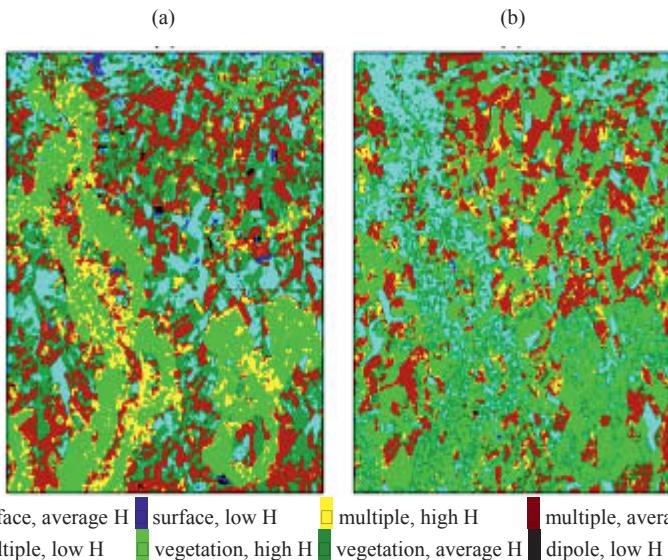


Figure 7.3. Gauss-Wishart multi-temporal ML classification over regions produced by radiometric segmentation (top) and Gauss-Wishart multi-temporal ML contextual (5×5 pixels) classification on data obtained using the linear vector MMSE estimator (bottom)



■ odd number of reflections ■ even number of reflections ■ diffuse backscattering ■ azimuthal asymmetry

Figure 7.5. Van Zyl classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data



■ surface, average H ■ surface, low H ■ multiple, high H ■ multiple, average H
 ■ multiple, low H ■ vegetation, high H ■ vegetation, average H ■ dipole, low H

Figure 7.6. Cloude and Pottier classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data

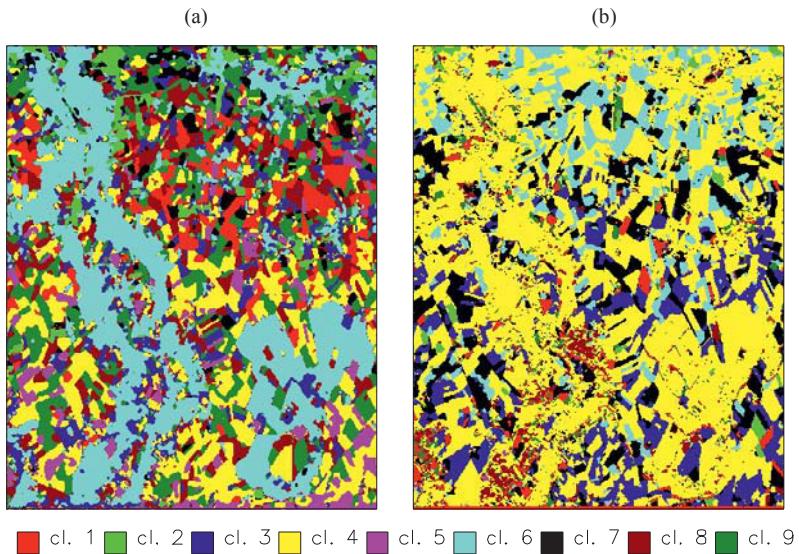


Figure 7.7. Rignot classification: (a) L-band; (b) C-band; Orgeval site case, AirSAR MAC-Europe '91 data

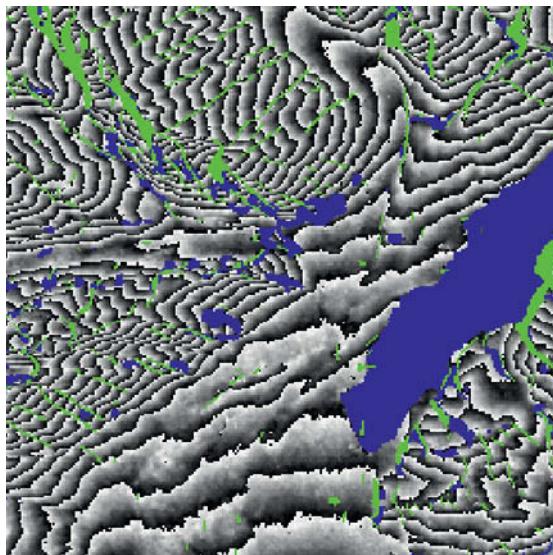


Figure 13.6f. Processing of an interferogram in order to prepare the phase unwrapping. The interferogram is filtered everywhere the phases are detectable; the non-unwrappable areas are marked: in blue for decorrelated areas, in green for discontinuous phase zones

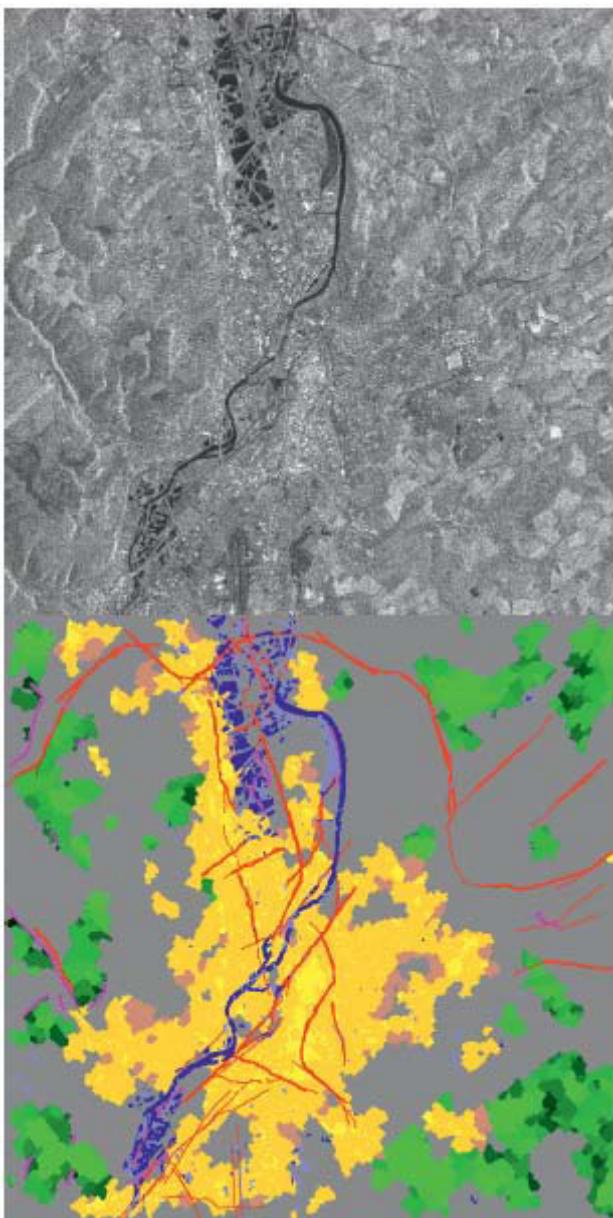


Figure 8.11. Illustration of network detection on an ERS-1 ©ESA image focused on the Metz area; the cartography set corresponds to the following legend – blue: rivers and lakes; red: road network; orange and brown: the decreasingly dense urban environment; green: forest; pink: relief peaks

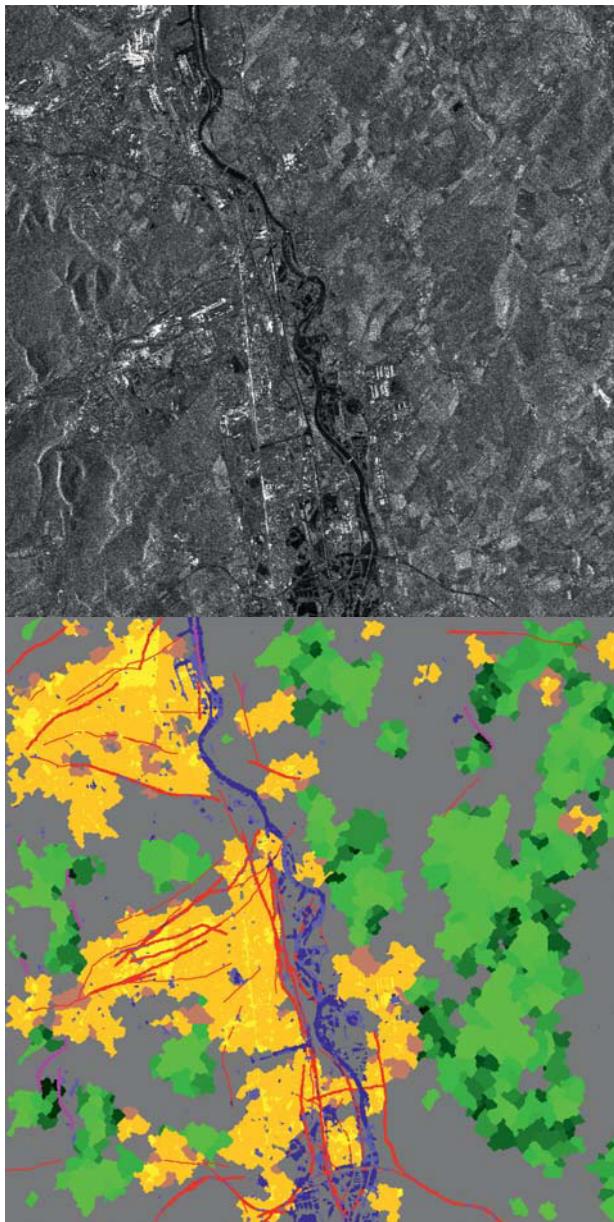


Figure 8.12. Illustration of network detection on an ERS-1 ©ESA image focused on the Strasbourg area; the cartography set corresponds to the following legend – blue: rivers and lakes; red: road network; orange and brown: the decreasingly dense urban environment; green: forest; pink: relief peaks



Figure 8.13. Detection of the street network in an urban area (downtown San Francisco).
The original ERIM image (X band, 2.5 m resolution) and the result
of a Markovian road detection

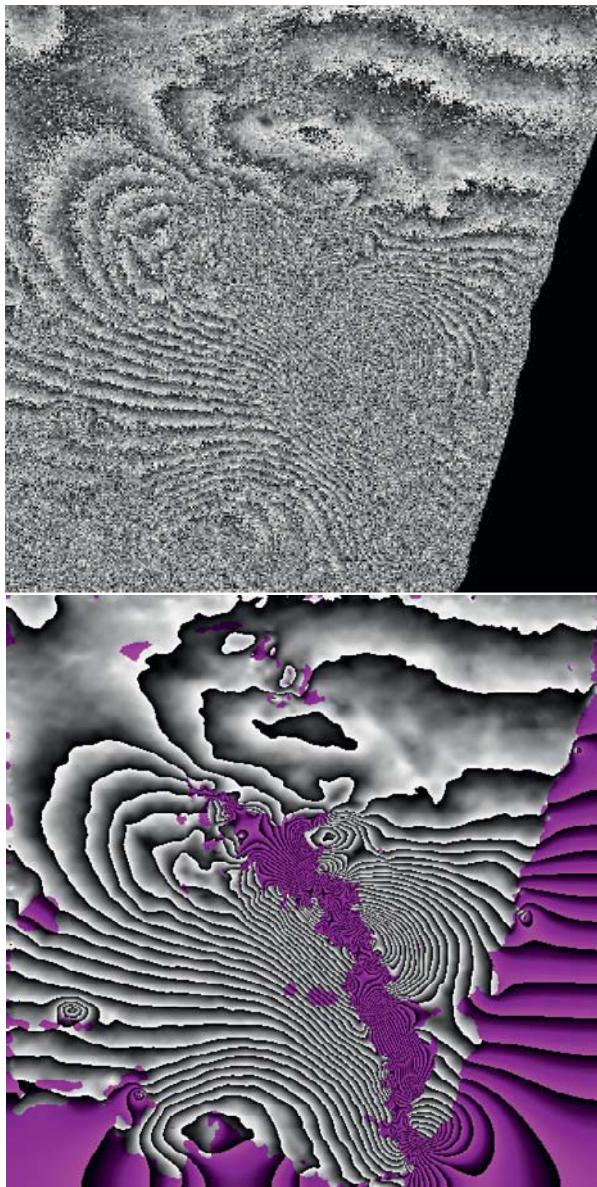


Figure 13.9. (a) 2,048 x 2,048-pixel differential interferogram on California's Landers fault © CNES; (b) phase unwrapped by the weighted least squares (wrapped for comparison). The areas in purple correspond to non-unwrappable phases