

Hierarchical Bayesian analysis of high complexity data for the inversion of metric InSAR in urban environments

Vom Fachbereich Elektrotechnik und Informatik der
Universität Siegen
zur Erklärung des akademischen Grades

**Doktor der Ingenieurwissenschaften
(Dr.-Ing.)**

genehmigte Dissertation

von

Dipl.-Phys. Marco Francesco Quartulli

1. Gutachter: Prof. Dr.-Ing. habil. O. Loffeld
2. Gutachter: Prof. Dr.-Ing. habil. M. Datcu
Vorsitzender: Prof. Dr.-Rer.Nat. W. Wiechert

Tag der mündlichen Prüfung:
18 März 2005

urn:nbn:de:hbz:467-1084

Abstract

In this thesis, structured hierarchical Bayesian models and estimators are considered for the analysis of multidimensional datasets representing high complexity phenomena.

The analysis is motivated by the problem of urban scene reconstruction and understanding from meter resolution InSAR data, observations of highly diverse, structured settlements through sophisticated, coherent radar based instruments from airborne or spaceborne platforms at distances of up to hundreds of kilometers from the scene.

Based on a Bayesian analysis framework, stochastic models are developed for both the original signals to be recovered (in this case, the original scene characteristics that are object of the analysis— 3D geometry, radiometry in terms of cover type) and the noisy acquisition instrument (a meter resolution SAR interferometer). The models are then combined to provide a consistent description of the acquisition process that can be inverted by the application of the so called Bayes' equation.

The developed models for both the scene and the acquisition system are splitted into a series of separated layers with likelihoods providing a probabilistic link between the different levels and with Maximum A Posteriori Bayesian inference as a basis for the estimation algorithms.

To discriminate between different Prior scene models and to provide the necessary ability to choose in a given set the most probable model for the data, a Bayesian model selection framework is considered.

In particular, a set of existing Gauss–Markov random field model–based algorithms for SAR and InSAR information extraction and denoising are extended by automated space–variant model–order selection capabilities whose performance is demonstrated by generating and validating model–complexity based classification maps of a set of test images as well as of real SAR data.

Based on that, a method for building recognition and reconstruction from InSAR data centered on Bayesian information extraction and data classification and fusion is developed. The system integrates signal based classes and user conjectures, and is demonstrated on input data ranging from on board Shuttle based observations of large urban centers to airborne data acquired at sub–metric resolutions on small rural ones.

To overcome the limitations of pixel based models and inference methods, a system based on stochastic geometry, decomposable object Gibbs fields and Monte Carlo Markov Chains

is developed and evaluated on sub-metric data acquired on both urban and industrial sites.

The developed algorithms are then extensively validated by integrating them in an image information mining system that enables the navigation and exploitation of large image archives based on a generic characterization of the data that is automatically generated.

Zusammenfassung

In dieser Dissertation werden strukturierte, hierarchische Bayes'sche Modelle und Schätzverfahren zur Analyse von komplexen mehrdimensionalen Fernerkundungsdaten vorgestellt.

Die entwickelten Methoden befassen sich mit der Problematik der Rekonstruktion und Interpretation von interferometrischen Radardaten mit einer Auflösung in der Größenordnung von einem Meter. Die betrachteten Daten beschreiben Stadtgebiete, wie sie von kohärenten luft- oder raumbasierten Sensoren aus großer Entfernung aufgenommen werden.

Basierend auf einem Bayes'schen Ansatz werden stochastische Modelle entwickelt sowohl für die Rekonstruktion der Szeneneigenschaften als auch für den verwendeten Sensor. Anschließend werden die Modelle kombiniert, um eine konsistente Beschreibung des Aufnahmeprozesses zu erreichen. Die entwickelten Modelle für die Szene und das Beobachtungssystem werden in mehrere getrennte Ebenen aufgeteilt. Dabei verbinden Wahrscheinlichkeiten die unterschiedlichen Ebenen. Die Basis für die Schätzverfahren liefert die Maximum A Posteriori Statistik.

Um zwischen unterschiedlichen A Priori Modellen der Szene zu unterscheiden und das Modell mit der höchsten Wahrscheinlichkeit auszuwählen, wird eine sog. Modellauswahl nach Bayes benutzt. Diese Methodik führt zur Entwicklung von einigen Algorithmen, die die Interpretation von Radar- und interferometrischen Radardaten von Stadtszenen erlauben. Im Besonderen werden einige bereits existierende Algorithmen zur Informationsgewinnung und Filterung von Radardaten, basierend auf Gauß-Markov-Zufallsfeldern, erweitert zur raumvarianten automatischen Bestimmung der Modellordnung. Die Leistungsstärke dieser Methoden wird durch modellordnungsbasierte Klassifikationen dargestellt.

Basierend auf diesem Wissen wird eine Methode zur Rekonstruktion von Gebäuden mittels interferometrischer Radardaten entwickelt. Die Methode integriert signal-basierte und nutzerrelevante Klassen durch Bayes'sche Informationsgewinnung, Fusion und Klassifikation. Die Leistung des Systems wird an Hand von raumbezogenen Fernerkundungsdaten gezeigt.

Um die Beschränkungen von pixelbasierten Modellen und statistischen Verfahren zu überwinden, wurde ein System auf der Grundlage von stochastischer Geometrie, aufteilbaren Gibbs-Objektfeldern und Monte Carlo Methoden entwickelt. Zur Evaluierung werden Fernerkundungsbilder verwendet, die große Städte und Industrieanlagen bedecken.

Die entwickelten Algorithmen werden anschliessend ausführlich evaluiert, indem sie in

ein Image Information Mining System integriert werden. Das System ermöglicht es, in dem Datenarchiv zu navigieren und es zu analysieren.

Acknowledgments

I am indebted to my supervisor at DLR Prof. Dr. Mihai Datcu for his scientific guidance and his encouragement throughout these years of work. I am also grateful to Prof. Dr. Otmar Loffeld of ZESS at the University of Siegen for giving me the opportunity to carry out this thesis. Further thanks go to Prof. Dr. Wolfgang Wiechert from the University of Siegen for being the chairman of the thesis committee.

An essential part in the development of the ideas presented in this thesis is due to the stimulating working environment I have encountered in DLR at the Institute for Remote Sensing Technology of Prof. Dr. Richard Bamler. In particular, the daily work and exchange of ideas and encouragement in the Image Science group were fundamental in providing focus and breadth to the everyday work: I would like to thank Herbert Daschiel, Andrea Pelizzari, Mariana Ciucu, Inez Gomez, Cyrille Maire and Patrick Heas. Thanks go to Vito Alberga, Miguel Lopez-Martinez and Jose-Luis Alvarez-Perez of the High-Frequency and Microwaves Institute at DLR.

Without the collaboration of colleagues and friends at Advanced Computer Systems in Rome this work would have had significantly less broad scope. Thanks go to Gastone Trevisiol, Roberto Medri, and the many others who supported my work and myself. In particular, people in the Image Mining group at ACS were fundamental in helping to aim at operational quality in the developed systems. Thanks go to Marco Pastori, Andrea Colapicchioni, Achille Valente, Annalisa Galoppo, Achille Natale and Massimo Manfredino for the encouragement, the advice, and the collaboration in the development, management and operation of the image archives used in the framework of the KIM project. Also, the collaboration of ETHZ was a key one in establishing a fruitful cooperation environment in which ideas could be developed and brought to an operational status. Thanks go in particular to Prof. Dr. Klaus Seidel.

For their friendly cooperations not only in professional matters I would also like to thank the many friends in München who divided with me the years of discovery and excitement: Sonia, Anna, Antonio, Silvia, Luis, Millah, Sandro, Daniela, Annette, Francesco, Leonardo and the so many others.

Finally, I want to thank Valeria Radicci and my family for their support during these years of studying.

Contents

Abstract	i
Zusammenfassung	iii
Acknowledgements	v
Contents	vi
List of Figures	x
List of Tables	xiv
1 Introduction	2
1.1 Advances in Bayesian inference for signal processing and analysis	2
1.2 On the information content of InSAR data at metric resolutions	4
1.3 Outline: the contribution of this thesis	7
I Preliminaries	10
2 Synthetic Aperture Radar Interferometry at meter resolution	12
2.1 Synthetic Aperture Radar	12
2.1.1 SAR radiometry and geometry	14
2.1.2 SAR processing and PSF	17
2.1.3 SAR statistics	23
2.2 SAR interferometry	25

2.2.1	InSAR principle and processing	26
2.2.2	InSAR statistics	29
2.3	The metric InSAR domain	30
2.3.1	Meter resolution SAR and InSAR for single isolated buildings	37
2.3.2	Metric SAR statistics in urban environments	38
2.3.3	High resolution urban SAR and InSAR inversion: state of the art	39
2.4	Summary	41
3	Bayesian modeling, estimation and decision theory for multidimensional signal analysis	42
3.1	Bayesian modeling and analysis	42
3.1.1	A principled approach to the composition of stochastic descriptions	43
3.1.2	Modeling interdependence in random variable sets: chain Rule and Markov property	44
3.2	Gibbs random fields	45
3.2.1	Lattices, sets, neighborhoods, cliques, sites, pixel values	45
3.2.2	Random fields in terms of potentials and energies	46
3.2.3	Markov random fields in terms of belief networks	46
3.2.4	Hierarchical models	47
3.2.5	Principle of inference for 2-level hierarchical models	48
3.2.6	Occam razor and Occam factor	49
3.2.7	From posterior distributions to estimates and decisions	51
3.3	Estimation and decision theory	51
3.3.1	Decisions as posterior ratios	52
3.3.2	Estimation and decision algorithms	52
3.3.2.1	Analytic derivation of MAP estimate	53
3.3.2.2	Exact inference by complete enumeration	53
3.3.2.3	The Iterated Conditional Modes algorithm	53
3.3.2.4	The Gibbs sampler and Monte Carlo Markov chains	53
3.4	Summary	54

II Hierarchical Bayesian modelling	56
4 Space-variant model selection in model-based information extraction	58
4.1 Model based image denoising and information extraction	59
4.2 Image models	60
4.2.1 Gauss-Markov random fields	60
4.2.2 Gaussian likelihoods for additive Gaussian noise	60
4.2.3 Gamma likelihoods for multiplicative fully developed speckle noise .	61
4.3 Locally-adaptive parameter estimation by evidence maximization	61
4.3.1 Fixed model order computational overview	62
4.3.2 Fixed model order processing examples and global model evidence computation	64
4.4 Model order selection	64
4.4.1 Variable order Gauss-Markov random field hierarchical modeling . .	68
4.4.2 Local evidence maximization in Gauss-Markov random field image models	69
4.4.3 Model order selection examples and comments	70
4.4.4 Spatially-variant model order selection for SAR data analysis	74
4.5 Summary	74
5 Information Fusion for scene reconstruction from Interferometric SAR Data in Urban Environments	76
5.1 The need for probabilistic modeling in scene reconstruction	77
5.2 Scene-to-data hierarchical models	78
5.2.1 Scene reconstruction Bayesian belief networks	78
5.2.2 User feedback and machine learning	80
5.2.3 Information extraction from backscatter and height image data . . .	81
5.3 Learning procedure	81
5.4 Urban scene reconstruction from InSAR data	83
5.4.1 E-SAR Dresden data analysis for settlement understanding	83
5.5 Summary	87
6 Stochastic geometrical modeling for urban scene understanding	

from a single SAR intensity image with meter resolution	88
6.1 Scene understanding from SAR: objects and their relations	89
6.2 Marked spatial point processes as prior scene models	91
6.2.1 Point and mark distributions	92
6.2.2 Modelling hierarchy	92
6.3 Scene posterior structure decomposition in Gibbs potential terms	94
6.4 The scene prior potential term	95
6.5 The data likelihood potential term	95
6.5.1 Extraction of strong targets from clutter in urban environments by inverse Gaussian statistics, decision theory–based classification . . .	95
6.5.1.1 Radiometric likelihood potential term	96
6.5.2 Geometric likelihood potential terms	97
6.6 Inference in Bayesian scene understanding with non–analytic posteriors . .	97
6.6.1 Monte Carlo Markov Chain Gibbs sampling of the scene posterior .	97
6.6.2 Scene configuration initialization	98
6.6.3 Scene configuration sampling and Gibbs dynamics	98
6.7 Overview of estimation method	100
6.8 Marked Point Process based scene understanding example	102
6.9 Summary	102
 III Evaluation and validation of results	 104
6.10 Content–based retrieval, information mining, scene understanding	106
6.10.1 Information extraction for application–independent data characteri- zation	106
6.10.2 The knowledge-based information mining system	107
6.10.3 Knowledge-based information mining evaluation results	109
6.11 Case studies with local model order selection in model based information extraction	112
6.12 Urban scene reconstruction from InSAR data	112
6.12.1 Urban land-use mapping from SRTM data	116
6.12.2 Large building recognition from SRTM data	120

6.12.3 Building recognition from Intermap data	120
6.13 Marked point process model-based building reconstruction from metric airborne SAR	124
6.13.1 Industrial scene understanding from metric SAR backscatter	124
6.13.2 Composed building understanding from sub-metric SAR backscatter	127
6.14 Summary	130
Conclusions	131
Summary of the obtained results	132
Evaluation of the obtained results and outlook	133
A The K-means clustering algorithm	134
B Non-analytical optimization and MAP estimation by simulated annealing	135
Bibliography	137

List of Figures

1.1	Growing spaceborne SAR sensor resolution with time, together with most relevant features for image understanding.	4
1.2	Scene geometry and SAR backscatter intensity for different sensor resolutions.	5
1.3	Scene understanding, resolution and reconstruction levels	9
2.1	SAR image coordinate system with sub-metric resolution SAR intensity image example.	13
2.2	SAR acquisition geometry for side looking configurations.	15
2.3	Slant and ground range.	16
2.4	Chirp signal characteristics.	20
2.5	Spectral characteristics of SAR data.	21
2.6	Theoretical SAR Point Spread Function amplitude.	22
2.7	Measured sub-metric SAR Point Spread Function amplitude.	23
2.8	SAR backscatter intensity autocorrelation.	24
2.9	Measured SAR image distribution examples.	24
2.10	Interferometric SAR geometry.	26
2.11	Interferometric SAR data processing principle illustration.	28
2.12	Absolute and wrapped phase surfaces.	29
2.13	Analytical InSAR data distributions.	29
2.14	Resolution-dependent interferometric SAR phenomenology.	32
2.15	Specular and diffuse scattering.	34
2.16	Electromagnetic wave paths for direct reflections, double bouncings and triple bouncings on a dihedral.	36
2.17	Projection of a building into slant and ground.	38
3.1	Bayes' law illustration.	43

3.2	Markov property for random variables: random variables are depicted as circles. Two neighborhoods are represented. First-order cliques for neighborhood N1 are depicted by closed shapes enclosing the variables. The three cliques C1, C2 and C3 form the first-order neighborhood N1. The diagram describes a belief network encoding statistical independence relations: variable A is independent of D (no path exists between the two), while A is independent of B if C is known. Independence and conditional independence are respectively denoted by $\cdot \perp \cdot$ and $\cdot \perp \cdot \cdot$. The exploitation of conditional independence relations allows the simplification of high complexity probabilistic descriptions.	45
3.3	Hierarchical model illustration.	47
4.1	Clique structures for different Gauss–Markov random field model orders. . .	59
4.2	Flowchart of the evidence maximization algorithm used for texture parameter estimation in the model based denoising and information extraction system seen as an Expectation–Maximization algorithm.	61
4.3	Flowchart of the evidence maximization algorithm used for contextual texture parameter and model order estimation in the extended model based denoising and information extraction system seen as an Expectation–Maximization algorithm.	62
4.4	Synthetic and natural realizations of random fields.	63
4.5	Speckled versions of synthetic and natural instances of random fields.	65
4.6	Denoising and information extraction with fixed model order on synthetic data.	66
4.7	Denoising and information extraction with fixed model order on natural image data.	67
4.8	Clean and speckled multiple model (autobinomial and GMRF with order 2) test image.	70
4.9	Clean and speckled multiple model (GMRF with orders 2 and 5) test image.	71
4.10	Clean and speckled multiple model (Brodatz “cloud” and “straw”) test image.	72
4.11	Model order estimation for synthetic and natural mixed images.	73
4.12	E-SAR X band SAR backscatter amplitude image over Oberpfaffenhofen, Germany.	74
4.13	Despeckled image and estimated model order for SAR intensity image over Oberpfaffenhofen, Germany.	75
5.1	Learning scene semantics through hierarchical Bayesian Networks.	79
5.2	Scene reconstruction processing chain.	81

5.3	Model based SAR backscatter–InSAR derived height information extraction system.	82
5.4	E-SAR Dresden - Ground truth optical image, interferometric land use, SAR backscatter intensity and SAR texture norm.	84
5.5	E-SAR Dresden - Ground truth optical image and generated land use map.	85
5.6	E-SAR Dresden Semper Oper- Ground truth optical image and interferometric land use.	85
5.7	E-SAR Dresden Semper Oper- interferometric SAR backscatter coherence and interferometric phase.	86
5.8	E-SAR Dresden Semper Oper- land use map obtained by unsupervised classification and Bayesian classification/fusion.	86
6.1	Meter resolution interferometric SAR phenomenology.	90
6.2	Levels in the hierarchical scene model.	93
6.3	SAR phenomenology considered in the scene understanding process.	94
6.4	Gibbs sampler dynamics.	100
6.5	The complete marked point process MAP estimation system.	101
6.6	Intermap AeS-2 data on Trudering test area.	103
6.7	Image information mining data analysis hierarchy.	106
6.8	Scheme of principle of ingestion chain operations.	107
6.9	Hierarchical modeling of image content and user semantic.	108
6.10	KIM coverage of Mozambique.	109
6.11	X-SAR scene reconstruction by data classification and fusion results.	110
6.12	ERS data scene reconstruction by classification and fusion results.	111
6.13	Meter resolution information mining example ground truth and intensity data.	112
6.14	Meter resolution information mining example results.	113
6.15	Uncalibrated optical and RADARSAT-1 fine mode data over Las Vegas, USA.	114
6.16	Filtered RADARSAT-1 data with locally estimated number of looks.	115
6.17	Automatically estimated despeckled amplitude and local estimated model order for the Las Vegas RADARSAT-1 scene.	116
6.18	Scene reconstruction SRTM backscatter intensity image and DEM.	117
6.19	Scene reconstruction procedure for SRTM data.	118
6.20	Bayesian classification/fusion results and ground truth for SRTM data.	119
6.21	Full resolution intensity and DEM SRTM data.	121

6.22 Full resolution SRTM Bayesian classification/fusion results and ground truth map.	122
6.23 Full resolution scene reconstruction results and ground truth map.	123
6.24 Views of the rural scene reconstruction test area.	124
6.25 Rural scene reconstruction test data.	125
6.26 Rural test scene reconstruction results.	126
6.27 Industrial scene understanding data and results.	128
6.28 Building understanding data and results.	129
B.1 Simulated annealing scheme of principle.	136

List of Tables

2.1	Resolution and information content in SAR data.	30
2.2	Recognizable object structures with varying resolutions in SAR data.	31
3.1	The Gibbs Sampling algorithm.	54
6.1	Summary of SAR data ingested in the KIM system and the level of data analysis considered.	109
A.1	A simple version of the K-means clustering algorithm.	134
B.1	Simulated annealing algorithm description.	136

Chapter 1

Introduction

One of the goals of signal processing is the extraction of information from data that are noisy, vague or otherwise affected by uncertainty. In many situations additional complexity is introduced by the non-stationarity of the data both in the useful information and in its distortions. From information theory we know that the most compact encoding of data is given by the probabilistic model that describes it best. However, to find tractable models for high complexity data is not an easy task.

This thesis explains and proposes new methods to model high complexity data and to develop algorithms for parameter estimation for hierarchical Bayesian models, in the framework of image information mining and three dimensional structure analysis and reconstruction. Its goal is the derivation and real-world validation of a paradigm for the extraction, characterization and construction of features tied to a three dimensional spatial domain starting from incomplete, multiple-source data with high complexity. Emphasis is placed on the study of Bayesian hierarchical modeling methods and on the problematic of model selection in the framework of machine learning with models having hidden parameters.

The analysis is motivated by the problem of urban scene understanding from meter resolution InSAR data, in which data cubes are acquired by the observation of highly diverse, structured settlements through sophisticated, coherent radar based instruments from airborne or spaceborne platforms at distances of up to hundreds of kilometers from the scene. The thesis therefore includes an application of the obtained theoretical results to the accurate recovery of geometrical information describing built-up, urbanized areas observed through a very high resolution interferometric synthetic aperture radar apparatus.

1.1 Advances in Bayesian inference for signal processing and analysis

The techniques of hierarchical Bayesian modeling for feature extraction and model based measure from incomplete, noisy data have been established since a number of years in the signal processing literature. Applications have ranged from frequency analysis (Jaynes,

1987) to modelling feedback in human–computer interaction studies (Pavlovic et al., 1997). The seminal works by Besag (Besag, 1974; Besag et al., 1995) were instrumental in introducing the ideas of stochastic Markov modelling to the field of image processing (Winkler, 1995; Chellappa, 1985; Chellappa and Jain, 1993). Geman and Geman (1984) introduced the techniques of Gibbs modelling and sampling to the field. Their ideas were applied in providing solutions to a number of problems from noisy image data filtering (Besag, 1986) to content–based image retrieval (Flickner et al., 1995; Schröder et al., 2000b).

Bayesian analysis is often characterized by the fundamental role played in it by prior distributions. The usage of subjective ones has often been the ground for objections and controversies. Jeffreys (1939) and Jaynes (1986) layed the ground for the development of techniques that can be applied to generate objective prior descriptions starting from a set of objective constraints to the problem and from the principle of Maximum Entropy. Furthermore, Bayesian analysis can be used to perform a principled choice of a prior model in a class of available ones by employing a two–level description of the problem under analysis in which inference is performed both on the model parameters and on the models themselves. MacKay (1992) and O’Hagan (1994) showed how model selection naturally maps into the maximization of the Bayesian evidence across a model class.

A number of the principles and techniques of hierarchical Bayesian modelling and two–level Bayesian inference for the modelling and estimation of noisy, nonstationary 2D signals are summarized by Datcu et al. (1998) and Schröder et al. (1998a). Their works introduce the general problem of estimation theory in a Bayesian framework centering on the properties of 2D Gibbs random fields and on their role in estimation. The focus of Walessa and Datcu (2000) centers instead on the extraction of reliable estimates of the parameters of these models from noisy, nonstationary observations in a two–levels Bayesian approach. Gauss–Markov random fields are used to describe textured radar backscatter images corrupted by speckle noise. The described system performs an estimation of the texture parameters of the clean image. The order of the model that is used as a prior description of the data is not an object of the estimation, though, and is considered a fixed input parameter instead.

Further works concentrate on object–based stochastic geometry models for the analysis of the structure of image data. In particular, based on Ripley and Kelly (1977), Stoica et al. (2000), Imberty and Descombes (2002) and Ortner et al. (2002) focus on marked point processes whose attached label processes describe the characteristics of the elements in a scene. Gibbs fields can then be used to describe the interactions between the objects (Cressie, 1991).

The problem of conducting estimation from the often non–analytic posterior distributions is approached by considering Monte Carlo Markov Chain methods (Winkler, 1995). Estimation is conducted by Expectation Maximization methods (Dempster et al., 1977) in a number of image denoising and information extraction problems.

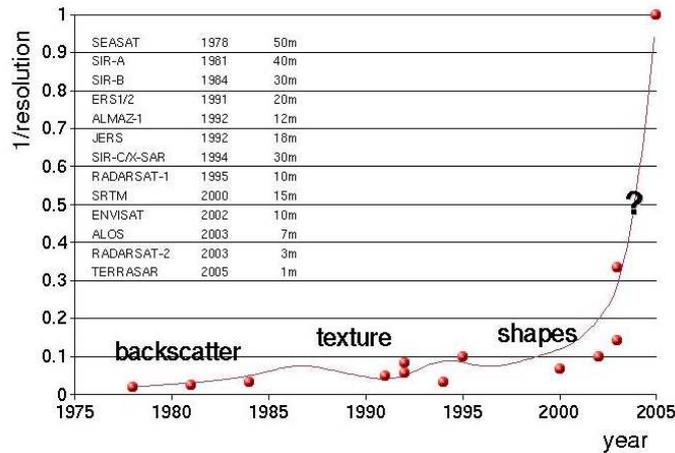


Figure 1.1: Growing spaceborne SAR sensor resolution with time, together with most relevant features for image understanding. This thesis addresses the area marked with a question mark on the upper right part of the plot, where relevant features are not yet established and therefore new methods are needed to fully exploit the information content in the data.

1.2 On the information content of InSAR data at metric resolutions

The reconstruction of buildings and of other typical settlement structures is a goal of great interest in an increasing number of applications related to the understanding and management of urbanized areas. At present, due to the broad availability of high resolution optical and of metric interferometric Synthetic Aperture Radar (SAR) data originating from both airborne and spaceborne sensors, new and more efficient methods of analysis are needed to address the challenges that are specific of the inherent complexity of the imaged scenes and of the acquisition systems themselves.

From the point of view of the traditional deterministic, pixel-based SAR and interferometric SAR data processing techniques, urban scenes are often regarded as problematic (D. L. Bickel, 1997):

- the scene: height discontinuities originated by vertical surfaces such as building walls are the norm in such data, while a wider dynamic range of the image shows up frequently, often saturating the receiver. Strong layover and shadowing effects imply an accentuated sparseness of the data with respect to the classical case. Multipath reflections and multiplicative noise from the sidelobes of the radar beam, as well as the presence of areas of low coherence are no more isolated deficiencies of generally simpler and more consistent data, but embody instead its most peculiar traits. Complex man-made objects and structures, moving elements such as cars or even varying housing details such as open and closed doors or windows show up at the imaging scales;

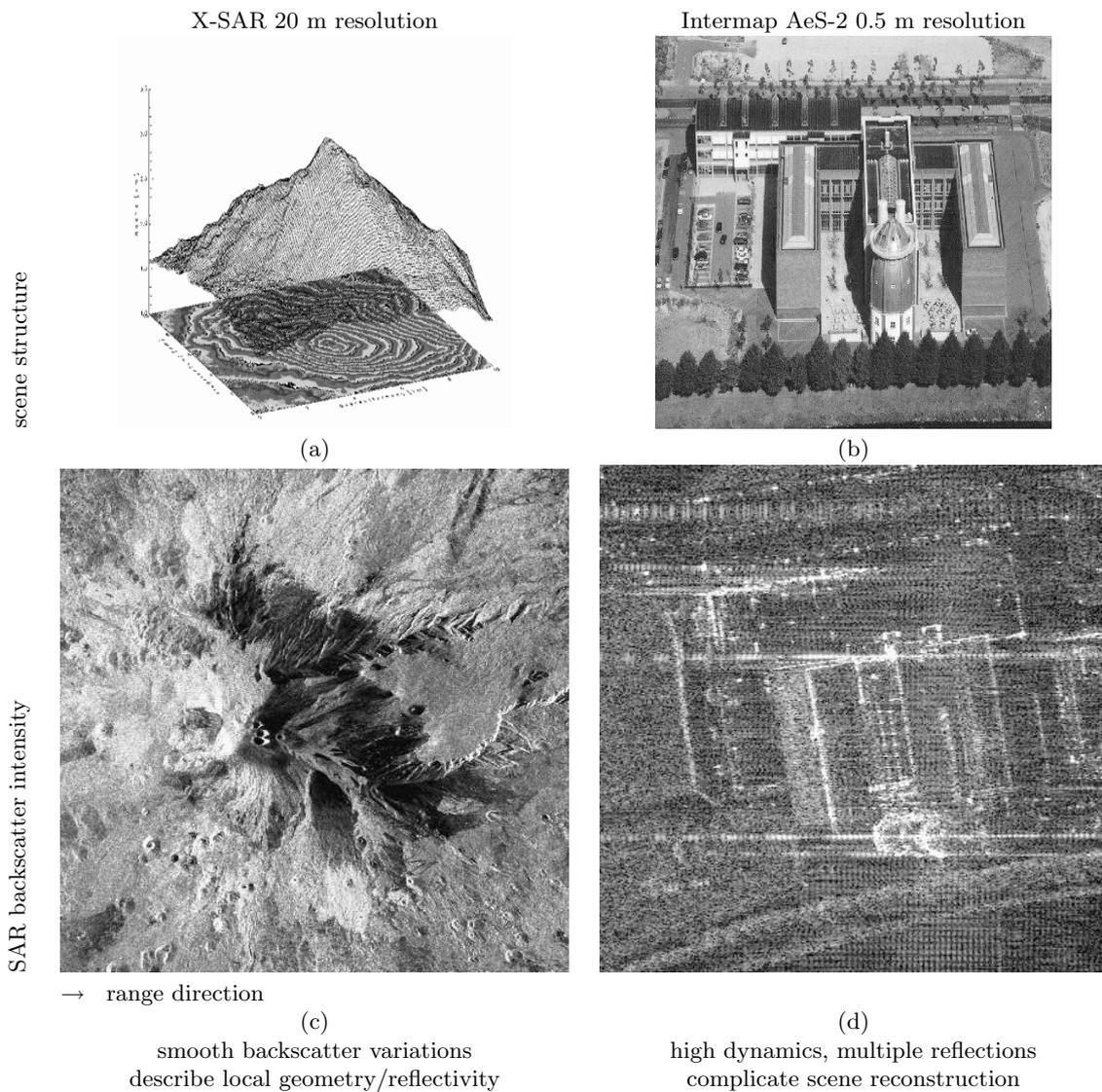


Figure 1.2: Scene geometry and SAR backscatter intensity for X-SAR view of Mount Aetna (a,c) at X band, 25 meters resolution, with the main volcanic cone visible (Schwäbisch, 1994) and Intermap AeS-2 view of Maastricht, the Netherlands, X band, 0.5 meters resolution, showing a single building composed by a number of square elements and by a round metallic tower generating multiple reflections clearly visible. With growing resolution, more complex models are needed to account for the mutated characteristics of observable scene element classes as well as for the peculiar effects introduced by sensors.

- the acquisition geometry: new corrections need to be taken into account for the attitude variations of the platform that carries the acquisition system and for the variations in the acquisition geometry from near to far range, that are much more pronounced than in the spaceborne case;
- the sensor: the high resolution of the acquisition system itself leads to image processing and SNR issues — speckle noise tends to change in nature, appearing in large correlated blobs, the sensitiveness to target texture changes completely and the whole statistics of the data must be re-evaluated. Strong isolated scatterers are mode common.

The acquired result changes so much in nature with respect to its traditional low-resolution counterpart on natural surfaces that the whole reconstruction approach needs to be shifted from the usual image processing based, two dimensional view to a full 3D domain in order to be able to explain and in the very end to exploit the effects and features that characterize this kind of data (Figure 1.2).

Significantly, up to now the problem of the reconstruction in the three dimensions of urban scenes from interferometric radar data has never been addressed in a unified, consistent framework by well-integrated algorithms and approaches: usually solutions originating from the well assessed problem of reconstructing low resolution imagery of natural surfaces are ported to the new operating environment and supplemented by manual intervention by the user (Heuel and Nevatia, 1995; Li et al., 1999) or are used as preliminary steps with moderate importance for some further processing carried out on higher resolution optical data that uses the obtained results as preliminary cues (Huertas et al., 1998). Strong assumptions are often made in the processing that tend to be appropriate for narrow application cases only (Gamba and Houshmand, 1999; Burkhart et al., 1996).

Unwrapped interferometric SAR phase surfaces are employed directly or after simple (e.g. morphologic) processing to obtain simple segmentation and classification maps that are used to initialize some shape from shade algorithm applied on coregistered optical imagery, or alternatively user assisted segmentation is followed by some back-projection operation in a shape from shade fashion (Bolter and Leberl, 2000).

While applications of stochastic modelling and analysis techniques have recently started to appear in the remote sensing field with particular reference to SAR with model based despeckling filters, knowledge based mining of large image databases and automatic target recognition and classification, the hierarchical modeling of data acquired on complex, detailed three dimensional structures in complicated environments has never been addressed: most of the studies up to now published limit themselves to the two dimensional case (Tupin et al., 1998; Schröder, 1999a) as a natural consequence of the simpler situation of natural environments imaged by lower resolution sensors. Traditional radargrammetric and interferometric techniques are used in the new environment along with ad-hoc methods that tend to provide very preliminary and partial results essentially by treating the peculiarities of the problem simply as shortcomings of the data that must be regularized in the processing instead of being used as additional sources of information.

1.3 Outline: the contribution of this thesis

The concept developed in this thesis is centered on the synergetic analysis of intensity and interferometric SAR data: a hierarchical model of the acquisition process and of its result is defined starting from a set of three-dimensional features that describe the scene and the sensor. The set is both deterministic and stochastic: while the deterministic section describes the SAR imaging geometry and its peculiar effects on the imaged scene and expresses a spatial description of the different scene structures, the stochastic part encapsulates instead prior knowledge about the SAR signal and details specific signal attributes.

The analysis is carried out as follows:

- a set of existing Gauss–Markov random field model–based algorithms for SAR and InSAR information extraction and denoising are extended by space–variant automated model–order selection capabilities (chapter 4) whose performance is demonstrated by generating and validating model complexity–based classification maps (figures 1.3a,b) of a set of test images as well as of real SAR data (Datcu and Quartulli, 2003);
- based on that, a method for building recognition and reconstruction from InSAR data centered on Bayesian information extraction and data classification and fusion is developed (chapter 5). Bayesian inference and networks are used to couple the different models specifying dependencies and to define further parameter estimation algorithms. The system integrates signal based classes and user conjectures, and is demonstrated on input data ranging from on–board Shuttle–based observations of large urban centers (figures 1.3c,d) to airborne data acquired at sub–metric resolutions on small rural center (Quartulli and Datcu, 2003b);
- to overcome the limitations of pixel–based models and inference methods, a system based on stochastic geometry, decomposable object Gibbs fields and Monte Carlo Markov Chains (Quartulli and Datcu, 2004) is developed. A semantic model is used as an important factor in the derivation of the results along with computer vision algorithms in order to provide a strong and robust discrimination criterion for the different elements showing up in the image as well as to simplify the separation of the different structures.

Such an approach (figures 1.3e,f) strongly reduces the need for the most problematic steps in the traditional interferometric SAR data processing chains (e.g. the phase unwrapping of the interferogram describing the city elevations model) and in the very end turn the sources of processing problems in standard interferometric methods into strengths in the form of additional sources of information providing added performance and robustness in a unified, coherent fashion (chapter 6);

- the developed approaches are validated on sub–metric data acquired on both urban and industrial sites (chapter III).

Such a conceptual shift, modeling the peculiarities of high resolution SAR acquisitions instead of considering them as artifacts to be treated by ad-hoc techniques generally separated

from the main corpus of the processing, produces as a final effect a significant simplification of the whole processing chain.

The methods here proposed are new with regards to both the estimation theory, with particular reference to the usage of a hierarchical mixed model for the evaluation and measurement of spatial features tied to a three dimensional domain, and to practice: to the best of our knowledge, none of the paradigms up to date employed for city reconstruction from SAR data shares the 3D hierarchical Bayesian modeling scheme that we propose, and none of them exploits in a unified fashion the multiple sources of information in SAR data to derive results in an attempt to overcome the limitations induced by the peculiar effects of the acquisition systems and of the complexity of the imaged scene.

The developed algorithms tend to construct a hierarchy of data models that leads from a set of data features to model-based descriptors and finally to features of the scene that has generated the data. While the first of the developed algorithms only considers the starting step of the modelling (from the data to the model-based features), all the others lead to the definition of scene characteristics in terms of scene elements or directly in terms of parametric, interacting scene objects.

This scheme provides results that, with respect to the reconstruction of geometrical features at least, outperforms the existing algorithms for urban area reconstruction from SAR data by overcoming their limitations related to the ad-hoc methods and conceptually separated processing steps and by turning into strengths the weaknesses of the existing approaches in the urban environment.

This work proceeds as follows: Chapters 2 and 3 introduce respectively Synthetic Aperture Radar systems with a resolution of the order of magnitude of one meter and the principles and techniques of Hierarchical Bayesian Modelling used in the remainder of the thesis. In part II, a set of techniques for feature extraction and scene understanding from high resolution SAR and interferometric SAR are introduced and explained in Chapters 4 to 6 and evaluated in Chapter III. Finally, conclusions are drawn.

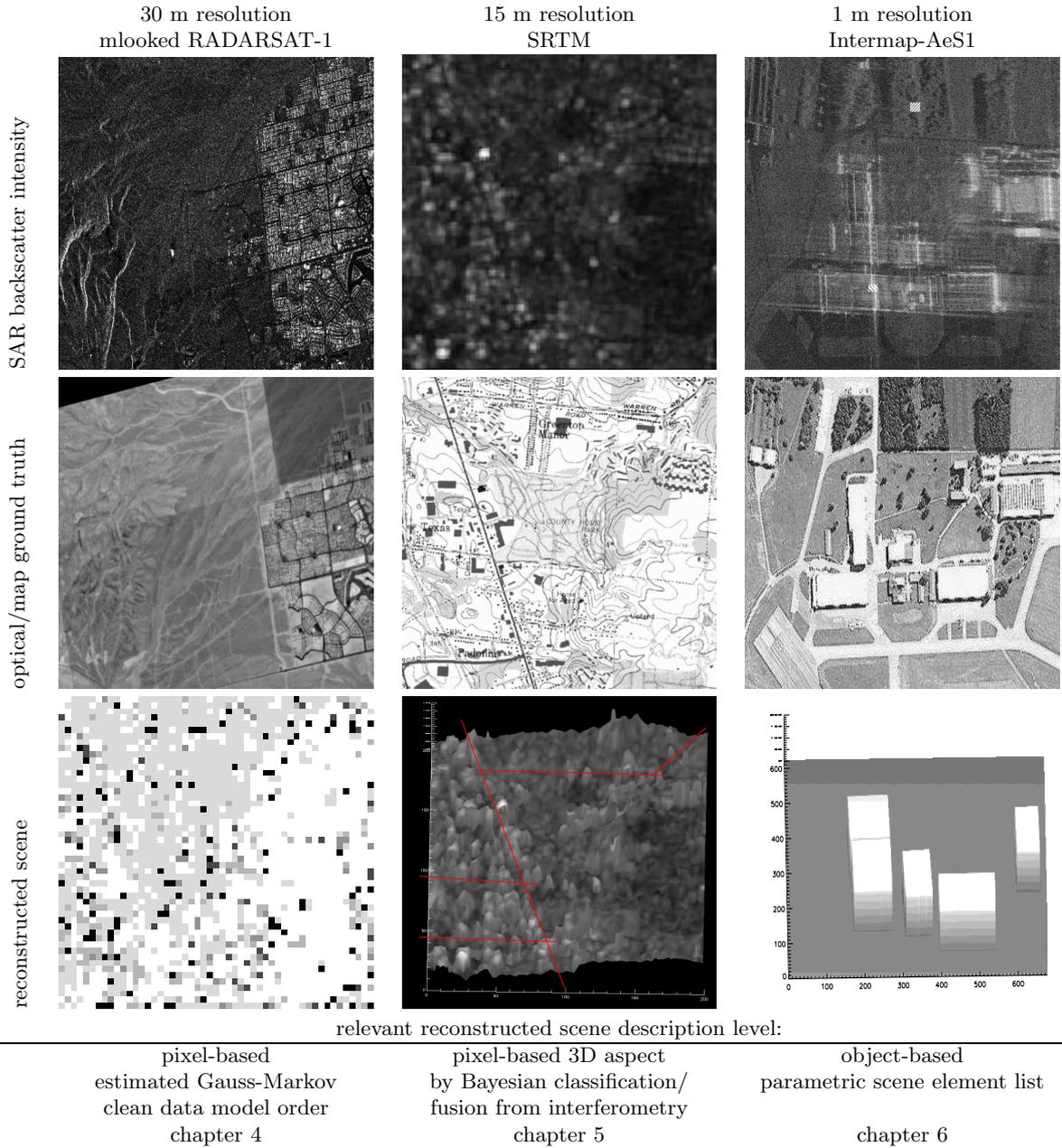


Figure 1.3: Scene understanding, sensor resolutions and obtained reconstruction levels. In the first row, SAR backscatter intensity of data acquired over urban areas by sensors at different resolutions (respectively RADARSAT-1, SRTM at X band, Intermap AeS-1). In the second row, either optical- or map-based ground truth for the datasets considered. In the third row, reconstructed scenes and obtained reconstruction levels. While for lower resolution data an essentially two-dimensional scene reconstruction in terms of land cover is usually sufficient (chapter 4) and can be extended to the third dimension by taking into account interferometric datasets and classifications in terms of typical scene elements (chapter 5), the information contained in high-resolution data is most naturally described in terms of scene objects in a three-dimensional domain (chapter 6).

Part I

Preliminaries

Abstract

In chapter 2, the characteristics of SAR and InSAR systems with metric resolutions are analyzed with reference to their performance in urban environments. Their geometric and radiometric properties are introduced together with the fundamental processing techniques used in their exploitation. The statistical properties of SAR and InSAR data are then detailed.

After introducing the peculiar phenomenology of SAR and InSAR with metric resolution in urban environments, traditional and model-based inversion algorithms in the literature are evaluated and compared.

Motivated by the challenges implicit in the exploitation of such high—complexity data in chapter 3, Bayesian modelling and estimation techniques for the analysis of multidimensional fields are introduced. After presenting the properties of Gibbs–Markov fields, hierarchical Bayesian models are introduced.

The second level of Bayesian analysis for model selection is presented. Bayesian estimation and decision theories are introduced together with modern posterior optimization techniques based on expectation maximization and on the Gibbs sampler and Monte Carlo Markov chains.

Chapter 2

Synthetic Aperture Radar Interferometry at meter resolution

Abstract

The geometric and radiometric characteristics of SAR and InSAR systems and the relevant processing techniques are introduced.

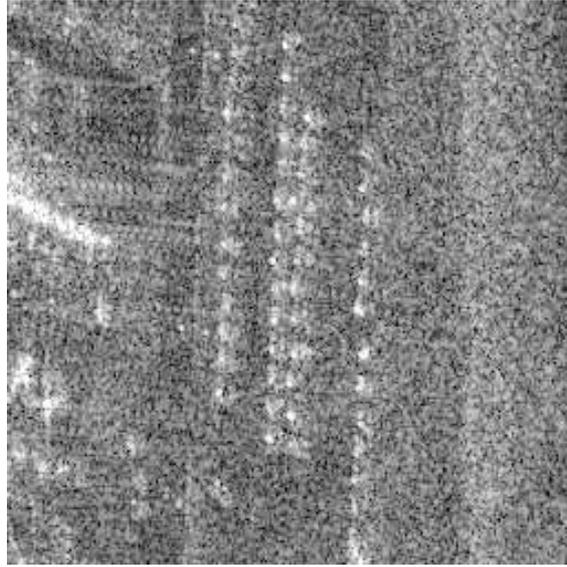
They are analyzed with reference to their performance at metric resolutions in urban environments: geometric effects such as layover, shadowing and occlusion are investigated together with the radiometric effect of smooth surfaces on speckle noise, backscatter texture and signal-to-noise ratio and on the appearance of multiple reflections of the incident radar beam. Strong isolated scatterer classes and behaviors are subsequently analyzed and related to the statistics of meter resolution data.

Traditional, simulation- and model-based inversion algorithms in the literature are evaluated and compared.

2.1 Synthetic Aperture Radar

The weather-resistant, all-time characteristics of the Synthetic Aperture Radar (SAR), as well as its peculiar sensitivity to scene characteristics such as dielectric properties and both large and small scale geometry (Curlander and McDonough, 1992), have made it the instrument of choice in a number of remote-sensing applications.

The instrument works by moving along its trajectory while recording echoes of a coherent modulated transmitted signal and by correlating them with a reference function that takes into account both the characteristics of the original signal as well as the expected rates of variation of these characteristics with both the along and across range sensor-target distance by considering the sensor platform motion characteristics.



↑ azimuth → range

Figure 2.1: SAR image coordinate system: unless otherwise stated, SAR images are presented with the range direction in the left–right direction, while the azimuth direction is represented as growing bottom to top. In the figure, view of a car parking at X band, resolution 0.5 meters: a number of isolated scatterers as well as large areas of very low SNR — very low energy return — are clearly visible.

The SAR principle The system can be described as a way of incrementing the resolution in the along track direction (the azimuth) by generating a synthetic virtual aperture that is orders of magnitude more extended than the physical one by coherently summing the echoes from a target in response to a coherent signal generated on board by a closely controlled oscillator. The usage of a longer antenna allows a corresponding increase in the azimuth resolution Δx . If, furthermore, the simplified expression the diffraction limit with λ denoting the carrier wavelength, W the antenna size, η_a the angular aperture of the beam and R the sensor to target distance is considered

$$\Delta x_{\text{eff}} \simeq \frac{\lambda}{2W} \cdot R$$

$$W_{\text{eff}} \simeq \eta_a \cdot R = \frac{\lambda}{W} \cdot R$$

a figure can be obtained for the system resolution

$$\Delta x_{\text{eff}} \simeq \frac{\lambda R}{2\lambda/W R} = \frac{W}{2}$$

that is independent of the range distance R .

A second approach describes the system as increasing the azimuth resolution by estimating Doppler shifts generated in the recorded signal by the angular difference between

the instantaneous position of the scatterer and a reference “zero Doppler” plane. From this point of view, the resolution of the SAR system depends directly on the quality of the frequency analysis performed: the instantaneous Doppler shift f_D of the target, generated by the relative speed v between sensor and ground, can be written as

$$f_D \simeq -\frac{2}{\lambda} \frac{dR}{dt} = -\frac{2}{\lambda} \frac{v^2}{R} t = -\frac{2}{\lambda} \frac{v}{R} \cdot x$$

and therefore

$$\Delta x = \left(\frac{2\lambda}{2v} \right) \Delta f_D$$

and considering that the Doppler resolution is related to the duration of the illumination of a single target on the ground

$$\Delta f_D \simeq \frac{1}{T} = \frac{\eta_a}{v} R = \frac{R\lambda}{Wv}$$

an expression for the azimuth resolution can again be written as

$$\Delta x \simeq \frac{\lambda R}{2v} \frac{Lv}{\lambda R} = \frac{W}{2}$$

according to Curlander and McDonough (1992).

The key point that makes the SAR principle possible is the presence of a stable oscillator from which all the signals are synthesized: the SAR is similar to a holographic device in which the recorded signal needs to be compared to a reference signal, a copy of the transmitted signal, to be visualized.

2.1.1 SAR radiometry and geometry

The radar equation The “radar equation” characterizes the energy of a backscattered wave given the properties of the target object. In the case in which the emitting and the receiving antenna coincide, if P_t and P_r are the emitted and the received power, then for an unfocussed SAR

$$P_r = P_t \frac{\lambda^2}{4\pi} \sigma \frac{G_t(\theta_i)}{4\pi R^2} \frac{G_r(\theta_i)}{4\pi R^2} \quad (2.1)$$

being G_t and G_r the emission and the reception gains at incidence angle θ_i , λ the carrier wavelength, R the antenna to target distance, and σ the radar backscatter equivalent surface of the target

$$\sigma = 4\pi \lim_{R \rightarrow +\infty} R^2 \frac{A_r}{A_t}$$

with A_r and A_t the received and transmitted signal amplitudes. The radar backscatter expresses the ratio between the energy received and that backscattered by the target. It depends on the incidence angle, on the dielectric characteristics of the target and on the rugosity of its surface with respect to the incident wavelength. The backscatter coefficient σ_0 is defined as the radar backscatter per unit surface. The intensity observed in the radar image is proportional to σ_0 (Curlander and McDonough, 1992).

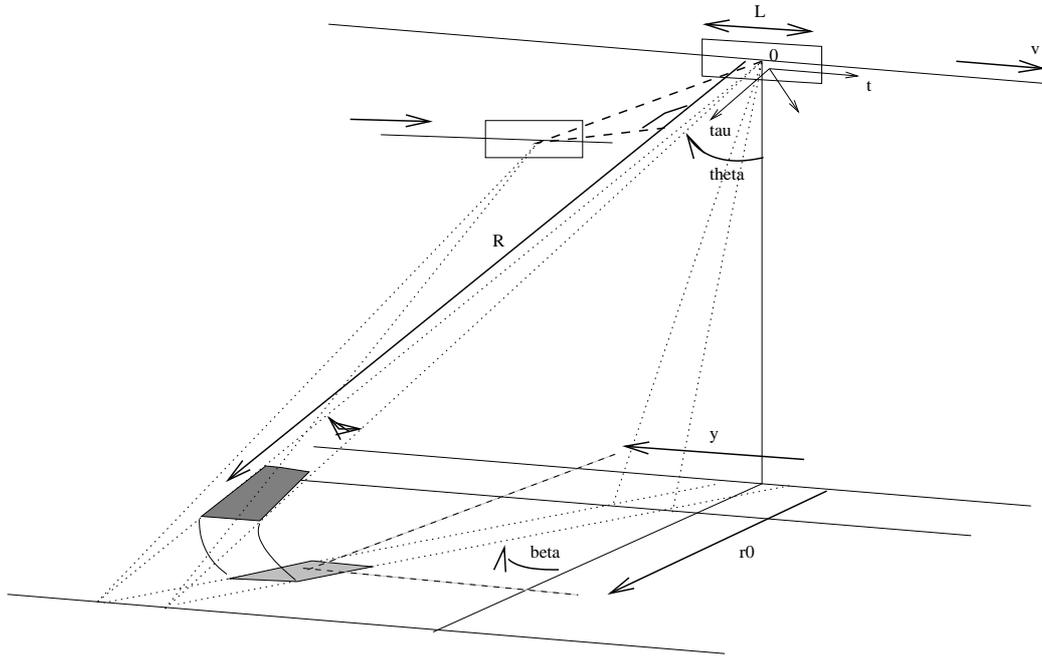


Figure 2.2: Acquisition geometry for side looking configurations: the asymmetry of the system with respect to the azimuth directions allows it to avoid the ambiguity right-left by the illumination of one side only of the trajectory.

SAR geometry To simplify the description of SAR systems and that of the data they produce, a coordinate system tied to the platform is employed. It has origin in the instantaneous position of the sensor, one axis x pointing in the instantaneous direction of motion of the sensor, the “azimuth”, and a second axis in the “slant range” direction, the vector linking the position of the antenna with the point nearest to it on the plane in which the illumination vector lies in the illuminated area on the ground.

Every point of a SAR image is then identified as belonging to a given point on the ground, to a specific equi-range surface (a sphere centered in the position of the sensor) and to an equi-Doppler surface (a double cone with vertex in the sensor position and axis coincident with the azimuth): each ground element can be identified from its range delay and its Doppler displacement in azimuth, and therefore the terrain can be described by a system made of concentric circles and coaxial hyperbole. The only potential ambiguity comes from the complete symmetry of the system with respect to the azimuth direction. To avoid this ambiguity, the antenna is not directly pointed to nadir but rather tilted laterally with respect to the the platform trajectory at a “look angle” θ_i in a so-called Side Looking geometry (see Figure 2.2) (Curlander and McDonough, 1992; Elachi, 1988).

Geometric effects The use of SAR images computed in the natural coordinates (slant range and azimuth) is complicated by the presence of geometric distortions intrinsic to the range imaging mode.

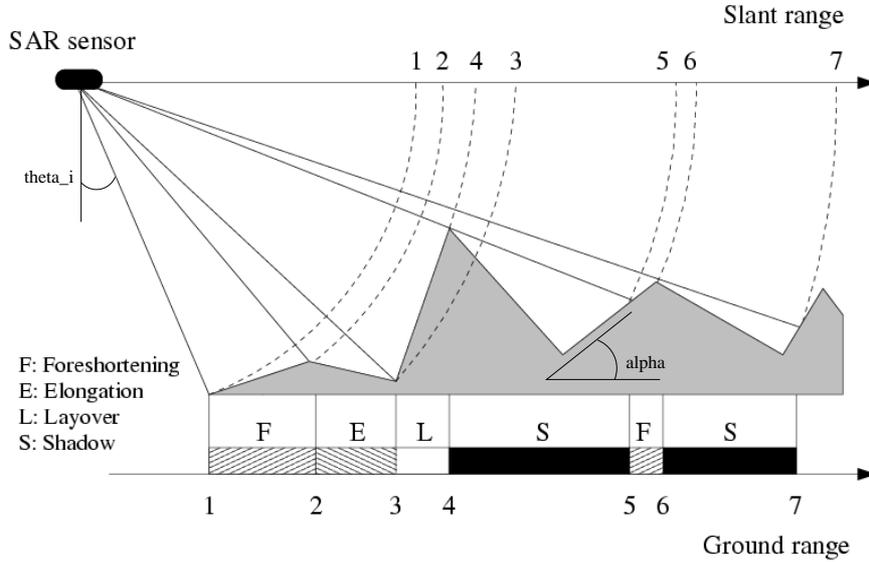


Figure 2.3: Slant and ground range (Walessa, 2000): the uniform sampling along the slant direction implies a dependence of the sampling step in ground on the local height of the illuminated terrain. The SAR system suffers from a series of prospective deformations that need to be compensated in the data processing step.

It is clear that if the system samples uniformly the terrain reflectivity in the slant and azimuth directions, it must sample the ground range direction with a density that depends on the terrain topography. The fact that the SAR system is set in a side looking geometry and the fact that it operates in cylindrical rather than in angular coordinates generates a number of geometrical effects.

Even if the illuminated area is planar, a constant resolution Δr in the slant range direction does not correspond to a similarly constant resolution, say Δy , on the ground range. In particular, the decrease of the incidence angle θ_i from near to far range leads to a decrease of the ground resolution

$$\Delta y = \frac{\Delta r}{\sin \theta_i}$$

and these results also apply to the ground range pixel dimension. For a surface slope of α , the resolution on the ground depends on the local incidence angle $\theta_i - \alpha$. Three cases are of interest (Franceschetti and Lanari, 1999):

- Foreshortening: $-\theta_i < \alpha < \theta_i$. It corresponds to a dilation or compression of the resolution cell on the ground with respect to the planar case, depending on the conditions $0 < \alpha < \theta_i$ or $-\theta_i < \alpha < 0$.
- Layover: $\alpha \geq \theta_i$. It causes an inversion of the image geometry. Ground elements with a steep slope commute with their bases in slant range, thus causing an extremely severe image distortion. A particular case is represented by the situation $\alpha = \theta_i$ corresponding to the compression of the area with this slope into a single pixel.

- Shadow: $\alpha \leq \theta_i - \pi/2$. In this case the region does not produce any backscattered signal.

Geocoding To generate SAR images with uniform and earth-fixed grids, a post-processing step is necessary; this is usually referred to as geocoding (Schreier, 1993). To perform it, knowledge of the location of each pixel of the SAR image with respect to a reference system is required. Processing of a single SAR data set generates a 2D SAR image related only to the two variables x, r of the cylindrical coordinate system (x, r, θ_i) . A solution to this problem is the use of SAR interferometry, that allows the determination of the further coordinate θ_i through the use of a second sensor. In cases in which InSAR data are severely distorted by foreshortening, layover and shadowing, and hence the traditional approach based on phase unwrapping tends to fail, the θ_i coordinate needs to be introduced into the system as external information in order to provide the full information needed for geolocalization.

2.1.2 SAR processing and PSF

The stop and go approximation Although in a somewhat physically inaccurate fashion, a fairly good description of the behavior of a SAR system in the along-track (or azimuth) direction can be derived in terms of a stop-and-go approximation as in Bamler and Schättler (1993). The antenna is supposed to be moving along a trajectory at constant altitude h with constant speed v . At well specified equally spaced positions along the trajectory, it stops, emits an electromagnetic wave, receives its echo and moves further along the trajectory.

The range chirp Focusing in the range direction is simply achieved by computing the correlation between transmitted wave and received signal. To obtain good resolution properties, signals with narrow autocorrelation and tendentially white spectrum have to be chosen. A “chirp” signal

$$g(\tau) = \exp(2\pi j k \frac{\tau^2}{2}) \text{rect}(\tau \frac{k}{B_\nu}) \quad (2.2)$$

can be used to obtain a complete analogy between the range and the azimuth directions. Since both the received signal and the transmitted chirp are long, the correlation is usually computed in the Fourier domain. Furthermore, modern SAR systems do not generate a chirp signal at each transmission act, but rather keep on board a digitized copy of the signal in order to better preserve its characteristics (see Figure 2.4).

The azimuth chirp An object placed on a flat ground surface below the system at range and azimuth coordinates r_0 and $x = v(t - t_0)$ has an overall instantaneous distance from the antenna of

$$R(r_0, t - t_0) = \sqrt{r_0^2 + v^2 \cdot (t - t_0)^2} \simeq r_0 + \frac{v^2(t - t_0)^2}{2r_0}$$

and a “range migration” of

$$\Delta R(r_0, t - t_0) = R(r_0, t - t_0) - r_0 \simeq \frac{v^2(t - t_0)^2}{2r_0}$$

where the last equalities hold if a parabolic approximation is considered. This range variation appears as a phase modulation $\hat{\varphi}(r_0, t - t_0)$

$$\exp[-j\hat{\varphi}(r_0, t - t_0)] = \exp\left[-j\frac{4\pi}{\lambda}R(r_0, t - t_0)\right] = \exp\left[-j\frac{4\pi}{\lambda}\left(r_0 - \frac{v^2(t - t_0)^2}{2r_0}\right)\right]$$

in the backscattered received signal. The term

$$\lim_{t \rightarrow t_0} \hat{\varphi}(r_0, t - t_0) = \left[-\frac{4\pi}{\lambda}R(r_0, t, t_0)\right] = -\frac{4\pi}{\lambda}r_0$$

is called “zero Doppler phase”.

This modulation results in an instantaneous “Doppler” frequency

$$f_D(r_0, t - t_0) = \frac{1}{2\pi} \frac{\partial}{\partial t} \hat{\varphi}(r_0, t - t_0) = -\frac{2}{\lambda} \frac{v^2}{r_0} (t - t_0)$$

and in a “Doppler” rate

$$FM(r_0) \equiv \frac{\partial}{\partial t} f_D(t - t_0) = -\frac{2}{\lambda} \frac{\partial^2}{\partial t^2} R(r_0, t - t_0) = -\frac{2}{\lambda} \frac{v^2 r_0^2}{R(t - t_0, r_0)^3} = -\frac{2}{\lambda} \frac{v^2}{r_0} < 0.$$

The instantaneous frequency at beam center $t_0 + t_C$

$$f_{DC} = f_D(r_0, t - t_0)|_{t=t_0+t_C} = -\frac{2}{\lambda} \frac{v^2}{R(r_0, t_C)} = FM(r_0) \cdot t_C$$

is called the “Doppler centroid” of the data. The energy of the data is distributed in azimuth around this center.

A signal characterized by a frequency increase like the one generated by the quadratic term in the last relations is again a chirp like the one considered in range.

SAR raw point scatterer response If a raw SAR point scatterer response (normalized to unity radar cross section) is considered

$$\hat{h}_a(\tau, t - t_0, r_0) = C(r_0) a_\beta(v(t - t_0 - t_C)/r_0) g(\tau - 2R(r_0, t - t_0)/c) \exp[-j\hat{\varphi}(r_0, t - t_0)]$$

in which the delayed range signal $g(\cdot)$ is multiplied by the instantaneous range-dependent modulation term $\exp[-j\hat{\varphi}(\cdot)]$ and by $C(\cdot)$, containing an R^{-2} amplitude range dependence term as well as an elevation antenna pattern, and an amplitude gain $a_\beta(\cdot)$ term function of azimuth time reflecting the shape of the two-way azimuth antenna pattern of the sensor, then if a unit δ -point is considered as a ground target having complex reflectivity

$$\hat{\gamma}^0(r, t) = \delta(t - t_0, r - r_0) \exp[j4\pi r_0/\lambda]$$

being $\delta(\cdot)$ a two-dimensional Dirac function and a response of

$$\hat{h}_a(\tau, t - t_0, r_0) \exp[j4\pi r_0/\lambda] = h_a(\tau - 2r_0/c, t - t_0, r_0)$$

then the ‘‘SAR data acquisition model’’ can be described in terms of a linear relationship between raw data $d(\tau, t)$ and object $\gamma^0(r, t)$

$$\begin{aligned} d(\tau, t) &= \int \int_{-\infty}^{+\infty} \gamma^0(r, t') h_a(\tau - 2r/c, t - t', r) dr dt' \\ &= \int_{-\infty}^{+\infty} \gamma^0(r, t) \star_t h_a(\tau - 2r/c, t, r) dr \end{aligned}$$

which is evidently a convolution in the azimuth dimension (denoted by \star_t) but is space variant in range.

If the range dependence of h_a is neglected (which is appropriate within a narrow swath), the last expression can be written as a two-dimensional convolution

$$d(\tau, r) \simeq \frac{c}{2} \gamma^0(\tau c/2, t) \star_r \star_t h_a(\tau, t, r_0) .$$

Focussed SAR point scatterer response The focussing in the azimuth direction is performed according to the theory of matched filtering by computing the correlation between the recorded radar echoes and the expected space-variant azimuth chirp.

The complex image $u(r, t)$ is then related to the raw data $d(\tau, t)$ via

$$u(r, t) = \int \int_{-\infty}^{+\infty} d(\tau, t') h_a^*(2r/c - \tau, t - t', r) d\tau dt' \quad (2.3)$$

which in the narrow swath approximation can be written as

$$u(\tau c/2, t) \simeq d(\tau, t) \star_r \star_t h_a^*(\tau, t, r_0) \propto d(\tau, t) \otimes_r \otimes_t h_a(\tau, t, r_0)$$

with \otimes denoting convolution.

The shape of the point scatterer response therefore obtained can be shown to be a cubic spline in azimuth and a sinc function in range:

$$s(r, t) \propto \text{spline}(2\nu/L \cdot t) \text{sinc}(2/c B_\nu r) \exp(j2\pi f_{DC} t)$$

being B_ν the available range bandwidth, with

$$\text{spline}(x) = \begin{cases} 2/3 - x^2 + |x|^3/2 & \text{for } |x| \leq 1 \\ 4/3 - 2|x| + x^2 - |x|^3/6 & \text{for } 1 < |x| \leq 2 \\ 0 & \text{else .} \end{cases}$$

The range and azimuth resolutions, defined as the half-power widths of $s(\cdot)$, are found as

$$\begin{aligned} \text{res}(r) &= 0.885 \cdot c/(2B_\nu) \Rightarrow \text{res}(y) = 0.885c/(2B_\nu \sin \theta_i) \\ \text{res}(t) &= 1.024 \cdot L/(2\nu) \Rightarrow \text{res}(x) = 1.024L/2 . \end{aligned}$$

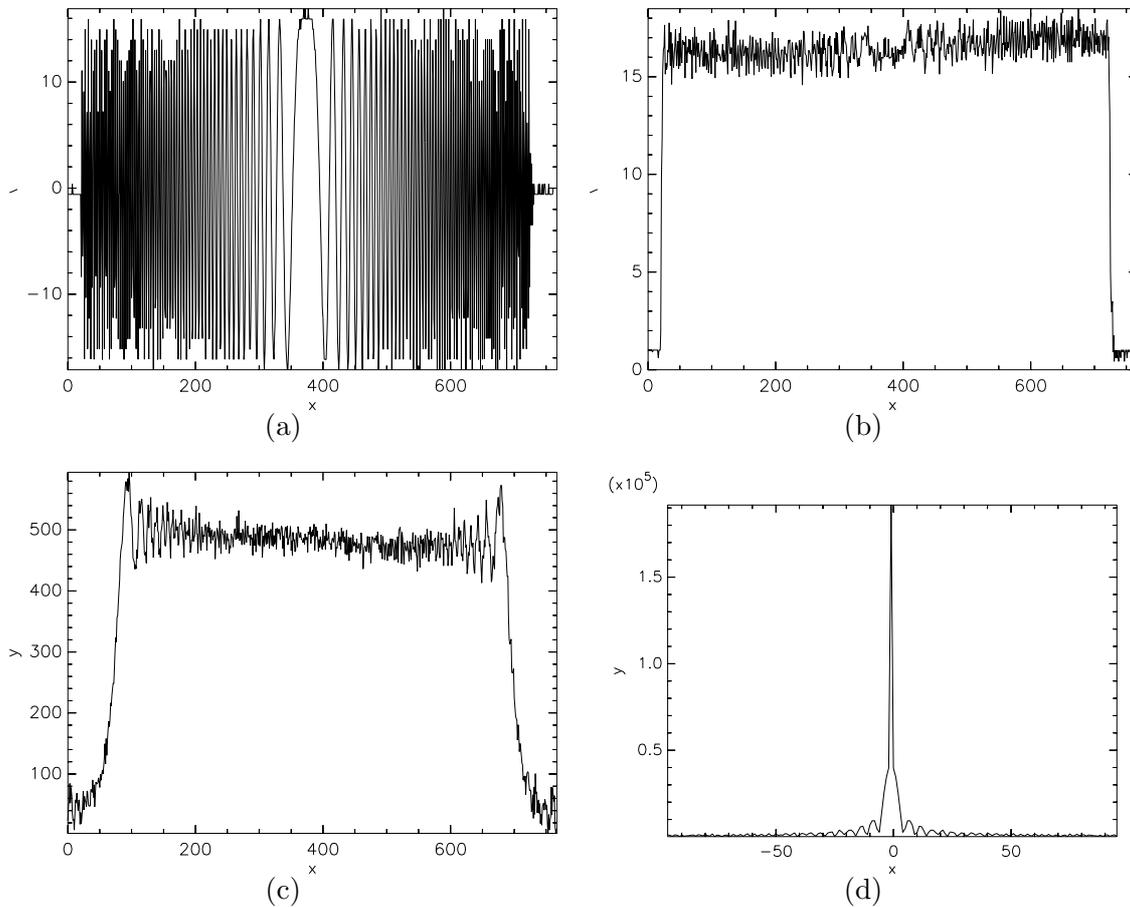


Figure 2.4: Real part (a) and amplitude (b), Fourier transform amplitude (c) and autocorrelation (d) of the 768 samples long ERS “chirp” signal. The characteristic linear frequency, constant amplitude, frequency modulation of the signal are visible in (a), (b) and (c). (d) shows the narrow autocorrelation of the signal.

The latter equation again states the fact that the azimuth resolution of a SAR system is in the order of half the physical antenna size L , irrespective of wavelength or sensor–target distance.

SAR processing strategies Although last equations look quite simple, the following peculiarities of the correlation kernel make SAR image formation a challenge for signal processing:

- the support of $h_a(\cdot)$ can be as large as a hundred range samples (due to range migration) and several thousand azimuth samples, which forbids the direct time domain implementation of equation 2.3 in most case;
- equation 2.3 is range-variant, i.e. an implementation via a two–dimensional FFT and

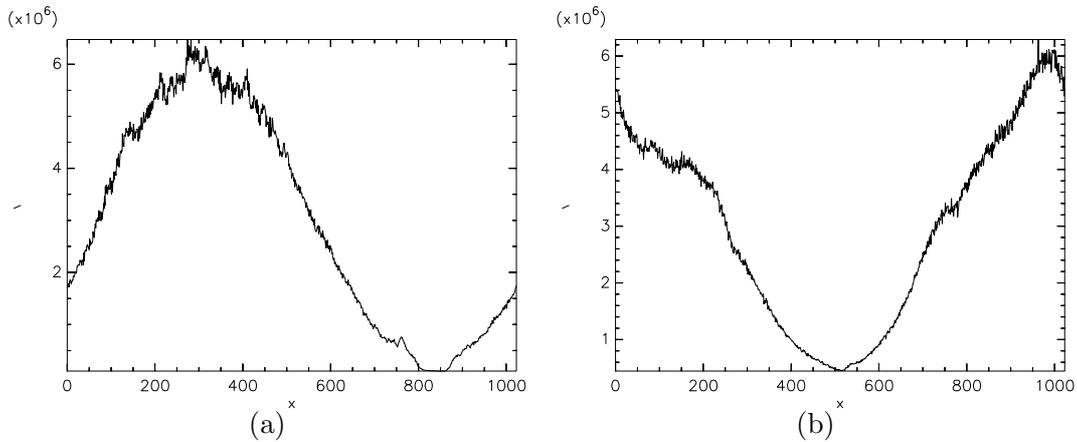


Figure 2.5: Azimuth (a) and range (b) spectra of sample ERS data. While the azimuth spectrum is Hamming weighted and shifted of a Doppler centroid frequency, the observed range spectrum is centered.

a single spectral filter multiply is only possible within a narrow range segment;

- due to range migration, the response function is inherently two-dimensional and non separable. Hence, the range-variance cannot be accounted for by simply using range dependent one-dimensional azimuth correlation kernels.

A direct implementation of the azimuth compression using a two-dimensional time domain correlation would be extremely computation time intensive. Therefore, frequency domain correlation methods are preferred. Common approaches to data focussing include:

- the range-Doppler approach (Cumming and Bennett, 1979; Wu et al., 1982; Jin and Wu, 1984), which operates in the range signal/azimuth frequency domain on the already range-compressed signal. Its main disadvantage is that the interpolation to convert the target trajectory along azimuth to a straight line given only by r_0 has to be carried out with a truncated kernel for efficiency purposes, which might cause image degradation. Furthermore, usually a single central frequency (usually the Doppler centroid, or a common mean one for the two antennas for physical interferometer configurations) is used for the entire processed block;
- wave-number techniques (Cafforio et al., 1991; Rocca et al., 1989; Bamler, 1992) make use of concepts from the field of wave propagation. The (already range compressed) signal is again treated in the two-dimensional frequency domain. The problem of the range variance of the two-dimensional compression filter can be overcome by a nonlinear mapping of the range frequency. This technique is sometimes referred to as ‘Stolt interpolation’ or ‘grid deformation’. The mapping can be split into a shift, a scale factor and a negligible higher order terms. It poses a trade-off between efficiency and image quality if phase preservation is a key issue;

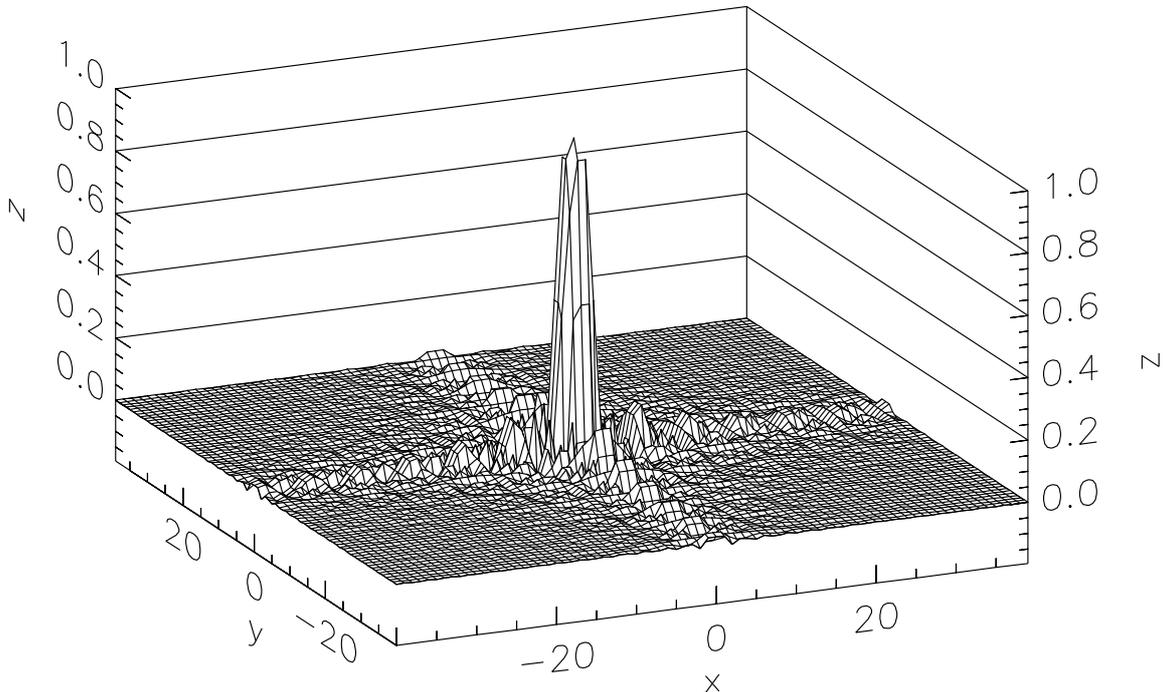


Figure 2.6: Theoretical SAR Point Spread Function amplitude.

- chirp-scaling algorithms (Runge and Bamler, 1992; Cumming et al., 1992; Raney et al., 1994) are based on the scaling properties of chirp signals. They avoid the use of interpolations (common to range-Doppler and wave-number techniques) during the processing. The range migration correction is performed by means of a scaling of the target trajectory in the range-Doppler domain: all the targets along the swath are shifted to have the same curvature as the one located at a reference range, usually at mid-swath;
- in an alternative approach, the need for range-dependent scaling of the along track pixel dimension is met by replacing the standard Fourier transform with a chirp z transform, the kernel of which includes a range-dependent correction (scaling) factor. The chirp z-transform can be computed by use of fast-Fourier-transform software, without need for zero-padding, instead of a post-processing interpolation step that can degrade either computational efficiency or accuracy (Lanari, 1995). Lanari and Fornaro (1997) subsequently showed that the chirp-z transform was actually a particular implementation of the chirp scaling algorithm. Both are equivalent to a scaled inverse Fourier transform (Loffeld et al., 1998).

It is important to notice the presence of sidelobes that may hinder the detection of weak main peaks in the vicinity of much stronger targets.

Resolution and multilooking The SNR of SAR data is often not sufficiently large for most remote sensing applications. The problem is usually handled by multilooking. It

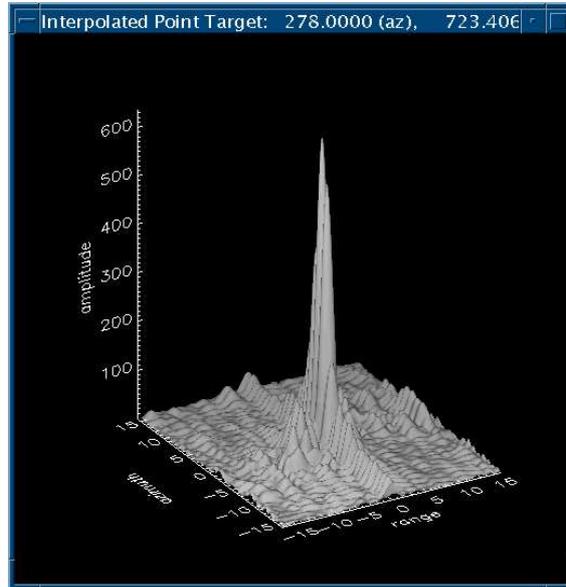


Figure 2.7: Measured SAR Point Spread Function amplitude at X band, 0.5 meters resolution.

consists of first dividing and then separately processing N non-overlapped portions of the SAR bandwidth. The incoherent average of the so obtained N SAR images improves the SNR by a factor N . However, antenna pattern spectral modulation, aliasing etc. render this improvement only an upper bound. Its effective value can be quantified in terms of an equivalent number $N' < N$ of uncorrelated samples; this number is usually referred to as equivalent number of looks (ENL).

Figure 2.8 represents the correlation coefficients computed in range and azimuth for a typical meter resolution SAR image at X band.

2.1.3 SAR statistics

Since the scattering properties of the illuminated scene can only be described in terms of statistical parameters, thus rendering the scattered field (and the SAR raw signal) a random process, SAR raw signals cannot be considered to be deterministic variables.

At resolutions of tens of meters, a SAR resolution cell is very large when compared to the centimeter wavelength of the illuminating electromagnetic wave. In addition, a large number of scatterers are generally present within each cell due to the roughness of the surface and/or the inhomogeneities of the scattering volume. The returned echo is the result of the coherent summation of all the returns due to the single scatterers: the phase of each single return is related to the distance between the sensor and the scatterer itself, their mutual orientation, and to the electromagnetic properties of the scattering material.

By describing the summation of the single vector responses as a random walk in the complex Gauss plane, it can be shown that magnitude and phase of the scatterers are

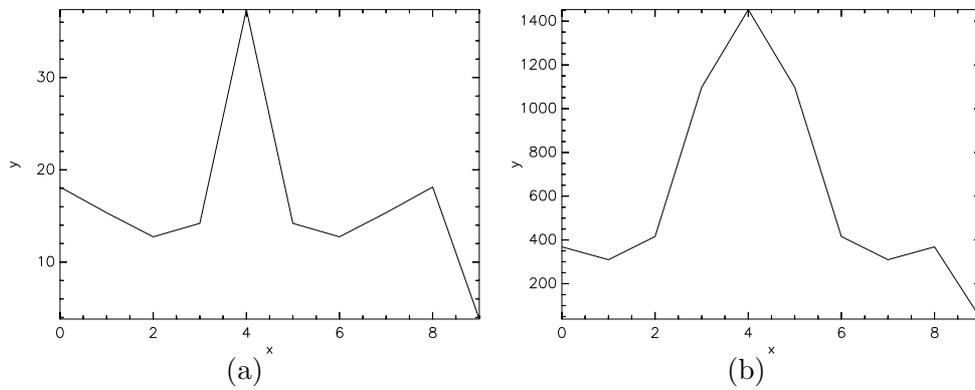


Figure 2.8: Measured SAR backscatter intensity autocorrelation values for high resolution SAR in range (a) and azimuth (b). Image was oversampled 2 times in azimuth.

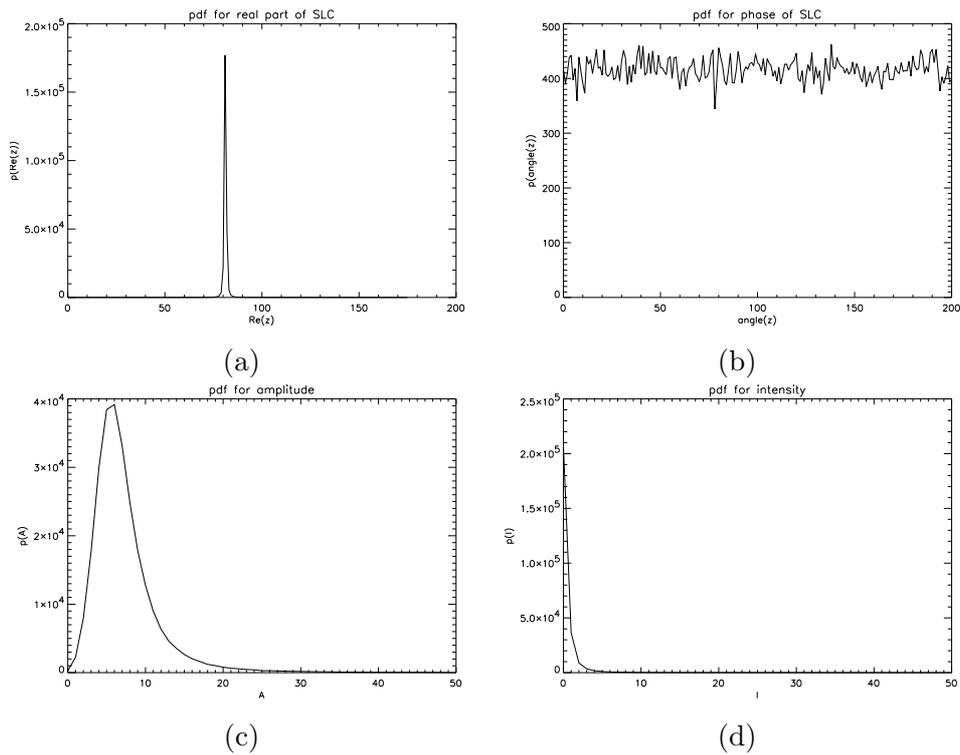


Figure 2.9: Measured SAR image distributions: (a) Gaussian distribution for real and imaginary part of complex SAR image, (b) uniform distribution SAR phase distribution (c) square-root Gamma distribution SAR amplitude distribution, (d) negative exponential for SAR backscatter intensity.

statistically independent. Furthermore, by simple derivations it is possible to show that if the number of individual wavelength-sized scatterers per resolution cell is high, real and imaginary parts of the received signal are Gaussian distributed with zero mean (according to the central limit theorem) and are statistically independent. The phase of the scatterers is uniformly distributed between 0 and 2π . Speckle is in this case assumed to be fully developed (Tur et al., 1982). This assumption does not apply for reflections from specular scatterers.

Speckle properties The characteristics of detected SAR images are quite different from those of both optical data and complex SAR data: in contrast to data acquired by incoherent systems, SAR backscatter intensity images appear to be affected by a granular and rather strong “speckle” noise (Goodman, 1975), an effect caused by random interferences between the electromagnetic waves reflected from the different scatterers present in the single resolution cell. Speckle becomes visible only in the detected amplitude or intensity signal: the complex signal by itself is distorted by thermal noise and signal processing induced effects only. Multilook intensity speckle adheres to a Gamma distribution characterized by the density

$$p(I = i_0, M_i = \mu_i) = \frac{L^L i_0^{L-1}}{\mu_i^L \Gamma(L)} \exp\left(-\frac{L i_0}{\mu_i}\right)$$

where L is the estimated number of looks of the data and μ_i its expected value.

Multi-look amplitude images are instead square-root Gamma distributed

$$p(A = a | M_a = \mu_a) = 2 \frac{a^{2L-1} L^L}{\mu_a^{2L} \Gamma(L)} \exp\left(-L \frac{a^2}{\mu_a^2}\right)$$

if μ_a is the square root of the expectation value of the amplitude.

2.2 SAR interferometry

varSynthetic Aperture Radar Interferometry (InSAR) is an extension of the radar concept that is made possible by the coherent nature of the signal (Prati et al., 1994; Massonnet, 1993). The contextual exploitation of data acquired on the same area from a number of slightly different positions allows the measurement of the local distance between the scene elements and the interferometer.

An indication of the geometric and radiometric stability of the scene is obtained by considering the interferometric coherence, an estimate of the amplitude of the normalized cross-correlation between the observations (Touzi, 1999).

The processing of SAR and InSAR data for the generation of backscatter intensity, Digital Elevation Model (DEM) and interferometric coherence measurements is a well-established field that has attained the operational stage since a number of years.

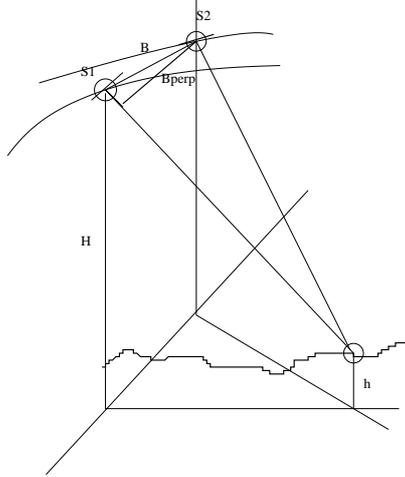


Figure 2.10: Interferometric SAR geometry.

2.2.1 InSAR principle and processing

If two antennas are involved, with 'baseline' spacing l across the range direction r , and if a point target is located in the plane orthogonal to the azimuth direction, at $(r = r', \theta)$, the signals collected by the antennas are

$$I_1 = |I_1| \exp\left[-j \frac{4\pi}{\lambda} r + j\varphi_{\text{scatter}}\right]$$

and

$$I_2 = |I_2| \exp\left[-j \frac{4\pi}{\lambda} (r + \delta r) + j\varphi_{\text{scatter}}\right]$$

if the effect of the random reflectivity term is neglected (Franceschetti and Lanari, 1999; Massonnet, 1993). From the two signals an interferometric pattern can be generated

$$\angle I_1 I_2^* = |I_1 I_2| \exp\left[j \frac{4\pi}{\lambda} \delta r'\right] = \exp[j\psi]$$

with \angle denoting the complex versor angle and

$$r + \delta r = r - l \sin(\theta - \beta)$$

and therefore

$$\psi = -\frac{4\pi}{\lambda} l \sin(\theta - \beta).$$

which relates the interferometric phase ψ to the cylindrical coordinate θ of the imaged point.

The transformation of slant range altitudes to ground range ones implies the possibility of obtaining height maps of the imaged areas solving the 3D location of the point because

all its three coordinates are determined. The third coordinate can be specified in terms of a length rather than an angle: if s is a slant range altitude,

$$\psi = 2 \frac{2\pi l_{\perp} s}{\lambda r_0}$$

and

$$s = -\epsilon \frac{\lambda r}{2\pi l} \psi \sim -\epsilon \frac{\lambda r}{2\pi l_{\perp}} \psi$$

where $\epsilon = 1/2$ for dual pass and $\epsilon = 1$ for single pass interferometry, respectively.

InSAR processing The first step in InSAR processing is the formation of the interferogram: complex images are coregistered (Carrasco et al., 1998; Scheiber et al., 1999) to sub-pixel precision by an interpolation whose factors are evaluated by local measures of signal correlation in the Fourier domain.

After the interferogram is computed by data common-band pre-filtering for the cancellation of non-common spectral contributions leading to decreased SNR in the interferogram (Gatelli et al., 1994), complex conjugate product of corresponding pixel values, phase extraction and phase pre-filtering techniques (Goldstein and Werner, 1997; Lee et al., 1998), multilooking and flat-earth phase removal algorithms (Carrasco et al., 1998) are applied to obtain a phase surface representing only the local surface heights.

Similar to intensity images, also in the interferometric case an average operation can be applied to reduce speckle effects and to improve the estimate of the interferometric phase. In this case the average step is carried out in the complex quantity $\hat{I}_1 \hat{I}_2^*$ and therefore is referred to as complex multilooking. This operation asymptotically ($N \rightarrow +\infty$) provides a maximum likelihood estimate of the phase interferogram (Curlander and McDonough, 1992; Franceschetti and Lanari, 1999).

Since the slant altitude is linearly proportional to the interferometric phase pattern, but the latter can only be measured in the $[-\pi, \pi[$ interval, appropriate phase unwrapping techniques must be implemented (Ghiglia and Romero, 1998; Costantini, 1998) to recover the full phase value φ (see figure 2.11).

Subsequent processing is necessary to convert the obtained absolute phase surface to a height map by multiplication with the local phase-cycle-height and to re-sample the resulting topographic map of the terrain in a map projection such as the Universal Transverse Mercator by geocoding techniques (Schreier, 1993) that allow the generation of ground range altitude maps.

InSAR coherence It is convenient for the statistical description of the interferogram to define the correlation coefficient

$$\chi = \left| \frac{E[I_1 I_2^*]}{\sqrt{E[I_1 I_1^*] E[I_2 I_2^*]}} \right|$$

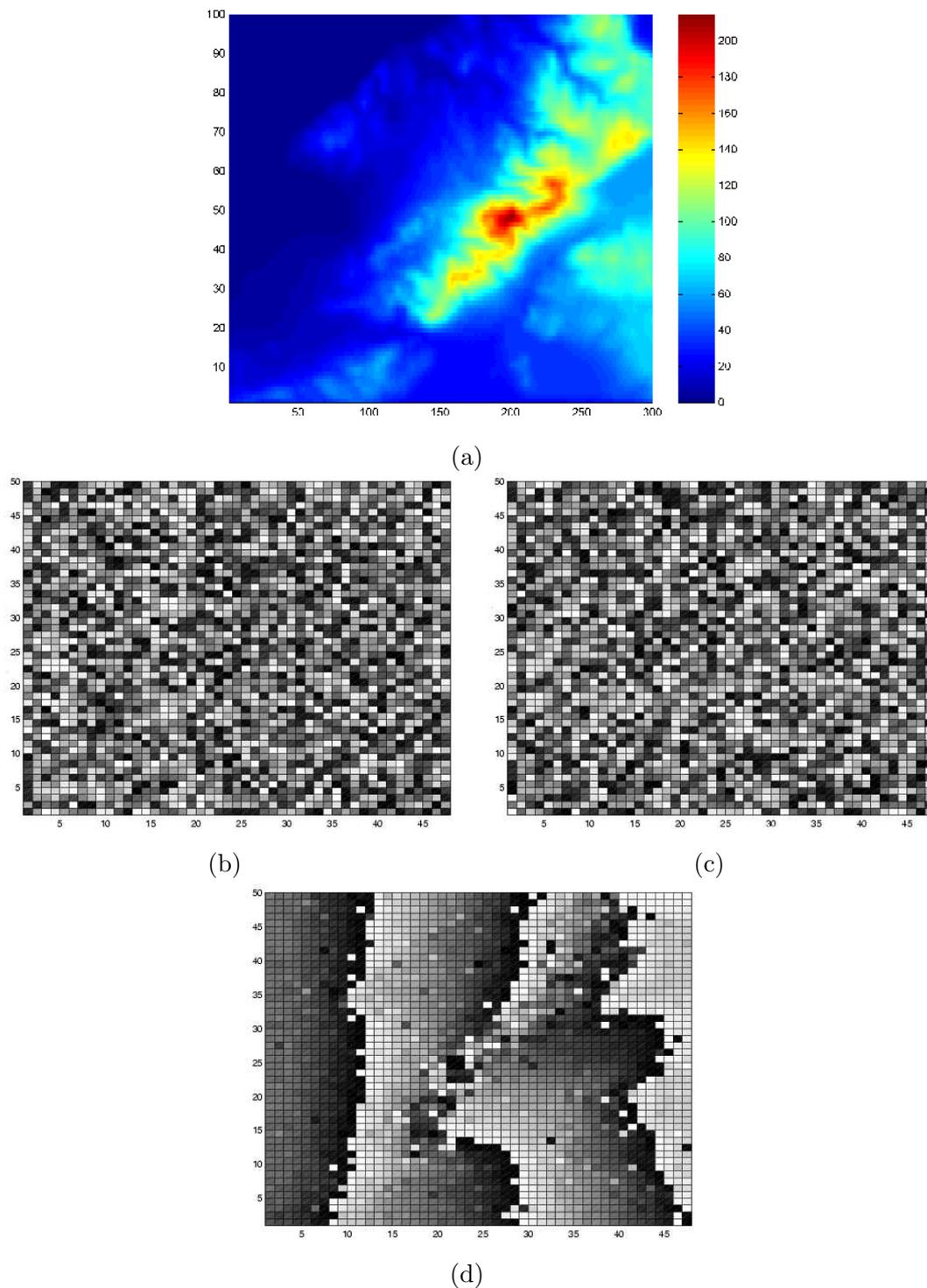


Figure 2.11: Interferometric SAR principle illustration: phase images (b) and (c) of relief (a) are combined into interferogram (d).

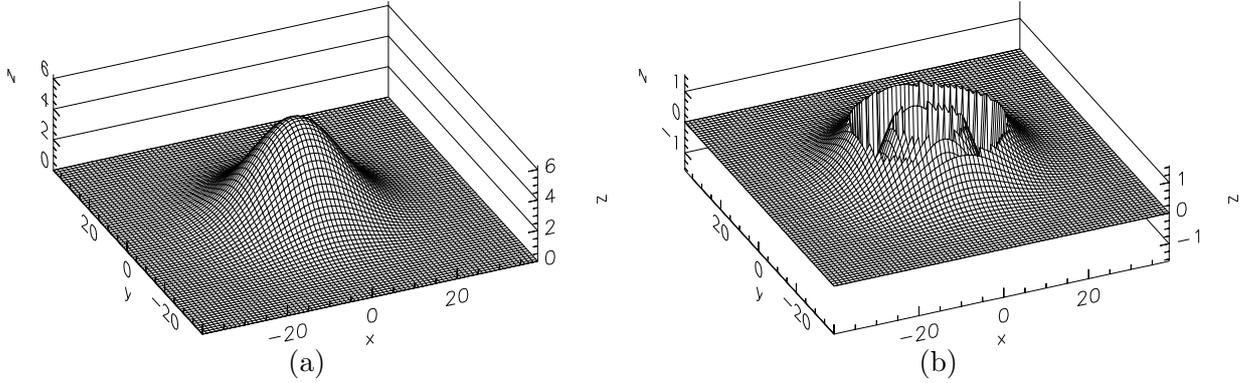
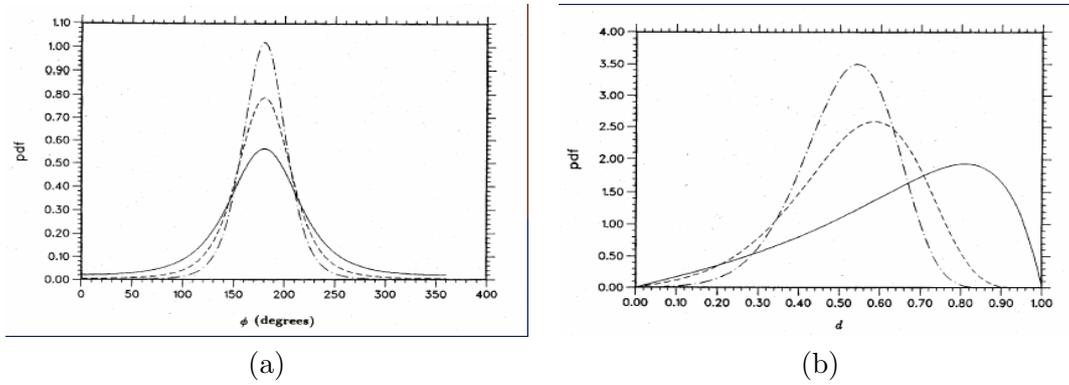


Figure 2.12: Absolute (a) and wrapped (b) Gaussian phase surfaces.

Figure 2.13: Analytical InSAR data distributions (Touzi and Lopes, 1996): (a) the Wishart distribution of interferometric phase noise for different looks: $L=3(-)$, $L=6(-)$, $L=10(-.)$ and (b) interferometric coherence distribution for different looks: $L=3(-)$, $L=10(-)$, $L=20(-.)$.

that has to be compensated for the topographic mean effective phase difference φ with $\varphi \in [-\pi, \pi[$ as in

$$\chi_0 = \left| \frac{E[I_1 I_2^*] \cdot \exp(-i\varphi)}{\sqrt{E[I_1 I_1^*] E[I_2 I_2^*]}} \right|.$$

The term χ_0 is usually referred to as interferometric coherence and provides an estimate of the local phase image SNR and of the small-scale geometric stability of the scene, among others.

2.2.2 InSAR statistics

The probability distribution function of the degree of coherence χ is derived by Touzi and Lopes (1996) to be

$$p(X = \chi) = 2(L-1)(1-\chi_0^2)^L \chi (1-\chi^2)^{L-2} F(L; L; 1; \chi^2 \chi_0^2)$$

for $L > 2$ and $\chi_0 \neq 1$, and being F the hypergeometric function.

Sensor resolution	Information sources	Applications for interpretation	Methodologies for data analysis
10-30 m	InSAR DEM, coherence, backscatter, some texture	natural scenes	Parameter retrieval, classification, DEM generation
2-10 m	Structure: - texture - linesedges -backscatter InSAR	natural scenes	
0.5-2 m	Strong targets	man-made scenes	analysis on sparse scattering

Table 2.1: Resolution and information content in SAR data.

The interferometric phase noise can instead be shown (Sarabandi, 1992; Lee et al., 1994; Touzi and Lopes, 1996) to be Wishart distributed: if φ is the true phase value and φ_n the associated noise, then

$$p(\Psi = \psi) = \frac{(1 - \chi^2)^L}{2\pi} \cdot F(L, 1; 1/2; \kappa^2) + \kappa \frac{\Gamma(L + 1/2)}{\Gamma(L)} \frac{(1 - \chi^2)^L}{\sqrt{4\pi} (1 - \kappa^2)^{L+1/2}}$$

being

$$\kappa = \chi \cos(\varphi - \varphi_n).$$

The Wishart distribution can be shown to correspond here to that expected for that of the phase of a sinusoid summed to complex, white Gaussian noise with $\text{SNR} = \chi^2 / (1 - \chi^2)$. A series of plots of different Wishart distribution functions with different values of the interferometric coherence are reported in Figure 2.13.

2.3 The metric InSAR domain

The phenomenology of urban high resolution SAR is peculiar and has to be studied specifically. With higher resolution, while natural targets show a typical fractal self-similarity property, man made scene elements, well confined in their scale of relevance, tend to change their appearance noticeably. A new level of complexity appears, determined by the presence in the scene of a number of objects whose influence on the data has to be understood. Historical buildings, moving vehicles, metallic structures need to be modeled and reconstructed.

Data geometric phenomenology at metric resolutions In the traditional approach to SAR data processing, a number of effects like layover or shadowing are regarded as data

Resolution	Object / Structure	Notes
2 m - 4 m	buildings, settlements forest, other vegetation	strong texture
10 m - 30 m	agriculture, industrial areas, vegetation	low texture
60 m - 100 m	drainage, geomorphology, agriculture	strong texture

Table 2.2: Recognizable object structures with varying resolutions in SAR data.

shortcomings that have to be regularized or filtered out in pre-processing steps. While this approach is justified in the case of natural surfaces imaged by lower resolution sensors, in the case of meter resolution SAR, and especially in urban environments, these hindrances become the norm rather than the exception and must be dealt with appropriately (Adragna et al., 2002):

- *shadowing* is a lack of signal caused by very steep surfaces in a way similar to that of geometrical optics: no energy is returned to the sensor in a given frame. The intensity images appear correspondingly dark, whereas interferometric coherence and phase turn to pure noise. Shadowing from buildings is ubiquitous in city scenes. It must be noted, though, that any triple scattering mechanism (figure 2.16c) corresponds to scattering from a virtual point whose position depends on the elevation of the building. This virtual point can easily be located inside the shadow area of a building. Therefore, in the case of meter resolution SAR data the presence of shadows from buildings is not guaranteed. Shadows, therefore, have only limited relevance when used as hints for the presence of buildings in an automatic scene understanding system;
- *foreshortening and specular reflections* represent an opposite lack respectively of range and of radiometric resolution to account for particular target geometries or radiometries (generally a combination of the two). The structure of the target becomes partially unobservable, and side-lobe effects might become important, especially in very high resolution data, thus further hindering information in other areas of the image. Specular reflections from house roofs and metallic structures appear frequently in cities;
- *layover* — an inversion of the range-to-height relation also caused by very steep surfaces — is the norm in scenes with buildings. Layover areas are generally simply masked out and ignored in the traditional processing approach;
- *occlusion* is a phenomenon also occurring in optical imaging. Although it is relatively infrequent in low resolution space-borne radar imaging, it becomes a key factor in the phenomenology of urban HR SAR imaging. Again, some information about the scene is missing from the data.

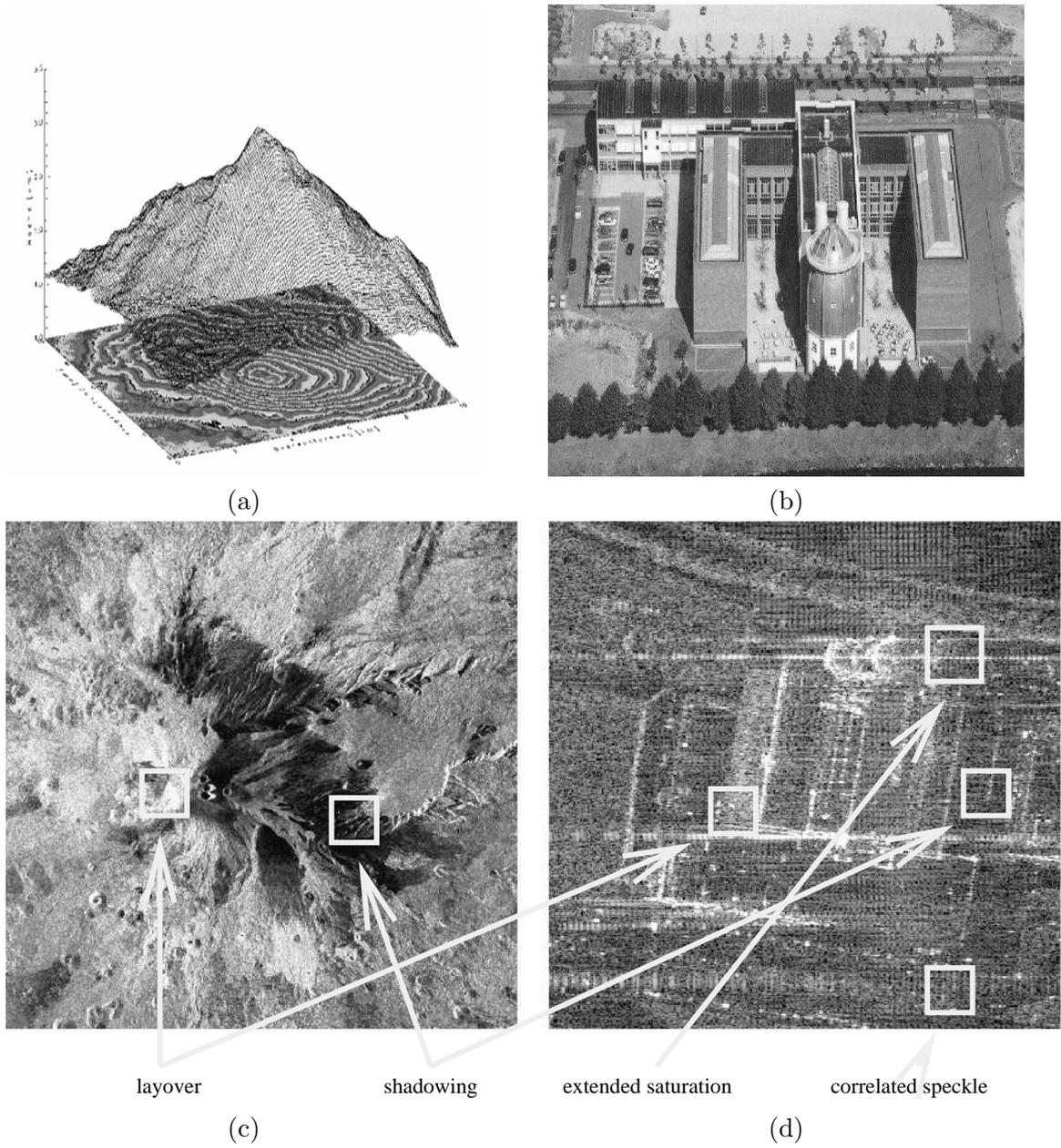


Figure 2.14: Resolution-dependent interferometric SAR phenomenology: scene (a,b) and basic signal features (c,d) for X-SAR image of Mount Etna (Schwäbisch, 1994) and for X band sub-metric resolution Intermap data over Maastricht, the Netherlands.

While all of these effects do appear in low-resolution images acquired on natural surfaces, their effect is usually limited to a small minority of the imaged areas (an important exception being alpine regions characterized by extremely steep slopes and very dense mountainous peaks). The affected image areas can, therefore, be left out of the processing results and integrated with data resulting from different sources. In high-resolution SAR data on urban environments, though, this exceptional cases become the norm, and the effects need to be dealt with appropriately. Furthermore, the fact that they are originated by the characteristics of the scene geometry means that they contain information on it: they should be consequently considered as information sources rather than hindrances. Further effects that become fundamental in metric urban SAR include:

- *along-range variations* of the acquisition geometry and therefore of the SAR PSF are an important factor in airborne acquisitions of SAR data: the incidence angle is very different at near and at far range, and therefore the sensitivity of the system to scene geometry can change noticeably. Reconstruction algorithms that operate on large datasets must be aware of the effect;
- *speckle* appears as a strong multiplicative noise that makes the radar cross section not directly observable. While in lower resolution SAR it is spatially uncorrelated, in HR data the fully developed assumption does not hold: noise appears in large correlated blobs in strict dependence with the scene geometry;
- backscatter *texture* is different than expected: much more textural information appears in the images, especially on non saturated, natural surfaces, but the statistics of noise (which now appears as correlated speckle) change in a significant way, imposing a redesign of analysis techniques; in urban environments, instead, the prevalence of specular reflections on smooth extended surfaces renders most of the scene as non textured;
- *low-return* areas are not limited to shadows: electromagnetic waves are returned with a definite strength from clear edges, while they tend to be scattered away from smooth surfaces, except in conditions of perpendicular incidence, when they saturate the receiver. In these conditions, returns from strong — e.g. dihedral — *scatterers* dominate the image extending well beyond the resolution cell;
- *isolated scatterers* are more frequent with increased resolution/wavelength ratio and with the presence of wavelength-sized man made metallic objects in the scene.

Scattering phenomenology Two main forms of scattering are possible for radar waves when they interact with a target surface:

- smooth surfaces tend to scatter the incoming way in a privileged direction: the scattering is specular;
- surfaces with high rugosities tend to re-radiate the received energy uniformly in all directions: the backscattering is diffuse.

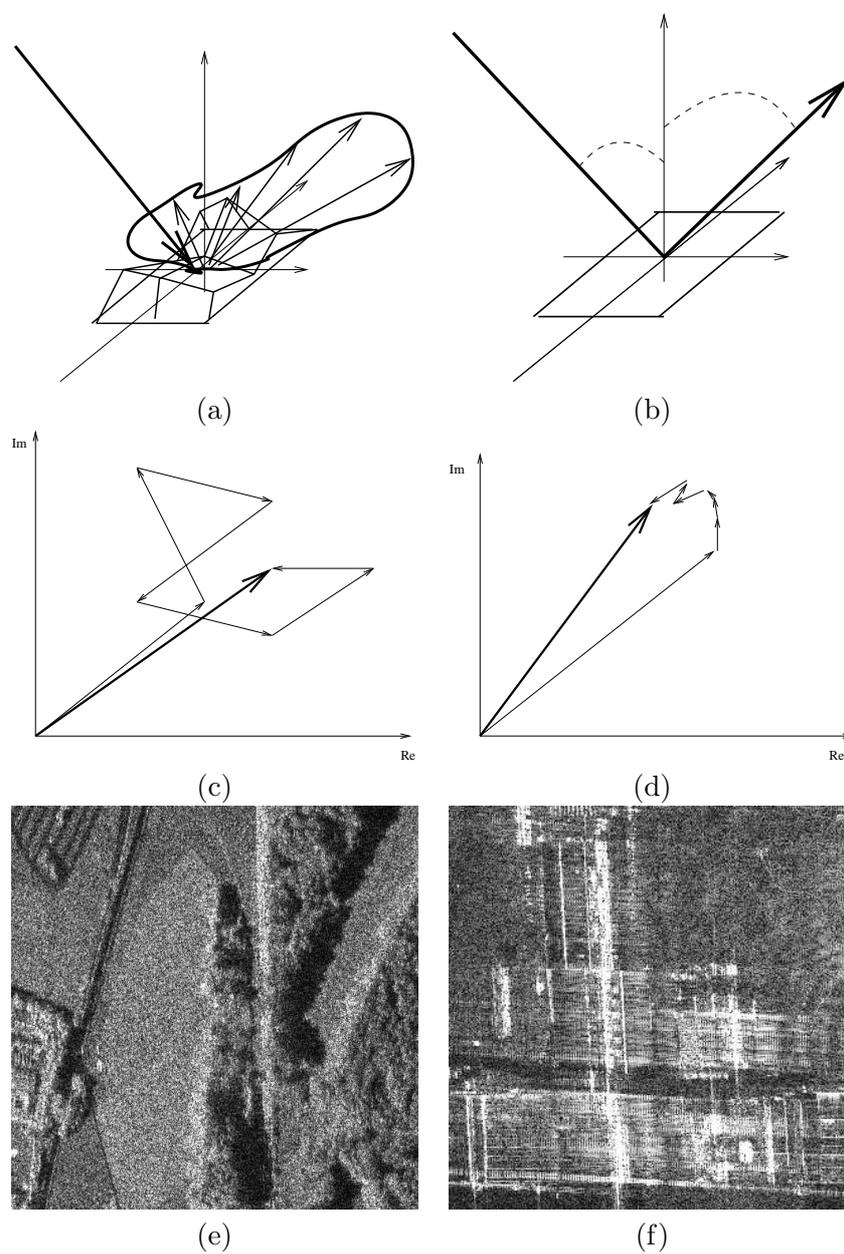


Figure 2.15: Scheme of principle: diffuse (a,c,e) and specular (b,d,f) scattering with resulting random walks in the complex Gauss plane.

According to the Rayleigh criterion, a surface tends to be smoother as the incidence angle increases and as the wavelength becomes larger, since any irregularities on the target become negligible. It must be noted that the Rayleigh criterion does not take into account the actual density of scatterers on the scattering surface. In the extreme case in which a single scatterer appears in a resolution cell (the probability increases as the wavelength and the resolution increase), the backscattered energy is dominated by the scatterer. Since the single scatterer tends to appear similar to itself when observed from different incidence angles in the azimuth direction (as well as in the range one), if its backscatter equivalent surface σ is sufficiently large it is possible to discriminate it by means of a correlation between different looks generated by considering different sub-apertures along the azimuth direction (Elachi, 1988).

Man-made targets, particularly such complex targets as those found in urban centers, contain scatterers with a variety of scattering behaviors (Rihaczeck and Hershkowitz, 1996):

- *discontinuities*: discontinuities are made of scatterers whose effective extents are relatively small in terms of the wavelength. Such discontinuities essentially act as fixed point scatterers, and their returns are usually relatively weak compared with those from other features on the target scene object. Discontinuities will become conspicuous on a target if the gain is set to high values or if the corner-like and cavity-like features on it are ineffective or inexistent;
- *smooth extended surfaces*: if a flat surface is illuminated at its broadside aspect, it will generate a huge return, commonly referred to as a specular return. Since the angular width η_a of the specular return is about equal to the ratio of the wavelength to twice the width of the antenna plate, this *specular reflection* is generated only within an extremely small angular section and is therefore rarely observed in practice. When it is, though, the return can be so strong that the other responses in the vicinity, much beyond the resolution cell in which the specular flash occurs, are more or less masked. In addition, for a practical radar the flash will generate inordinately high range and Doppler sidelobes, which tend to mask other responses.

When, on the other hand, smooth surfaces are observed out of the small broadside aspect, they tend to scatter most of the energy away from the incidence direction in a specular fashion. This makes them appear as areas of very low SNR in the data containing no usable information, often indistinguishable from shadows.

Since smooth extended surfaces are the norm in man-made objects such as vehicles, streets and many important classes of buildings, they determine a large part of the phenomenology of SAR data in urban environments;

- *corners*: an ideal trihedral, whose three sides are perpendicular to one another, acts as a triple bounce reflector and has the unique property that its effective phase center is at the corner point, independently of the aspect angle and of the frequency within very broad limits. The corner thus acts as a fixed, strong point reflector. A sizeable corner reflector that closely approaches the ideal trihedral generates returns that compete in strength with those from flat plates at perpendicular incidence.

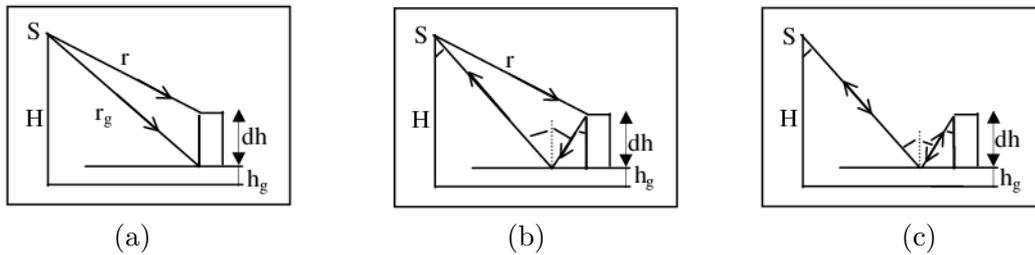


Figure 2.16: Electromagnetic wave paths (Simonetto et al., 1999) for direct reflections (a), double bouncings (b) and triple bouncings (c) on a dihedral.

Real corners only poorly approximate the ideal trihedral, but anyway tend to trap the incoming wave in a complicated fashion that usually cannot be explained by multiple-bounce reflections.

The general corner-type reflector typically gives much stronger returns than discontinuities: it will appear with a very bright phase center (the point from which the return appears to emanate, giving the effective position of the reflector as measured by radar).

- *multiple reflections*: with fixed azimuth and fixed target altitude, a number of target echoes will appear on a typical scene building image ordered by increasing distance or by decreasing altitude:
 - direct reflections on the borders of the structure edge;
 - direct reflections on the ground/structure edge;
 - double bouncings on the border, accumulating on the pixel corresponding to the edge: double-bounce propagation at the extended dihedral corner reflector between the ground and the building wall is mapped to the location at the building footprint. This leads to a line of bright intensity in azimuth direction at the edge of the building wall and ground;
 - higher-order bouncings.

Multiple bouncings generate an ambiguity in the interpretation of structure images and in the evaluation of its altitude, since every phenomenon corresponds to a different altitude. Their energy is normally much lower than that of direct bouncings on structure edges and of scattering on the structure base edge. They might be therefore difficult to observe.

Interferometric data phenomenology The processing of distance information acquired by interferometric techniques also becomes more complicated.

- *interferometric phase*, describing the distance between sensor and scene targets and hence the geometry of the scene, is often regarded as the key content in interferometric SAR observations. In metric resolution data over urban areas it cannot be directly

used as in traditional systems, though, unless many other information sources are recognized as such — instead of being considered as shortcomings in the data — to gain in robustness and discrimination power. Correct phase unwrapping is a complicated problem, and noise has to be accounted for without lowering the resolution if the characteristics of many of the scene elements have to be extracted;

- *interferometric coherence* images have different appearance: both geometry and radiometry tend to segment the scene into isolated patches of very coherent and very noisy responses. Simply increasing the size of the estimation windows loses geometrical information and raises further estimation problems.

The scene understanding problem is made more difficult by the compromises that are necessarily made in the design of very high resolution radar systems:

- spatial (and temporal) *baselines* might often be sub-optimal;
- *signal-to-noise* ratio issues might make the interpretation more difficult;
- *motion-compensation effects*, again for airborne acquisitions, can complicate the phenomenology: once the data is transformed to ground coordinates, strong deformations and even rotations of the PSF can be visible and need to be accounted for in inversion algorithms.

The increased resolution, on the other hand, means that contextual *geometric information* content is acquired at finer scales, and thus describes in greater detail every structure in the scene: backscatter, texture, edges and strong targets are evidently still valuable descriptors even in the new scenario. What is needed, though, is to relate them to a three-dimensional domain providing a natural context for their understanding.

2.3.1 Meter resolution SAR and InSAR for single isolated buildings

The reconstruction of man-made objects in dense urban areas from SAR imagery is complicated by inherent geometric constraints.

The incidence of phenomena like layover, shadow and multi-path signals (Schreier, 1993; Simonetto et al., 1999; Stilla et al., 2002) has to be considered with care in urban areas (see Figure 2.17).

If the building in Figure 2.17a is illuminated by a sensor located in F , in the slant range image the points A , B and C will appear ordered according to their range distance. Hence, the point A of the building footprint will appear as A' , behind point B' and in-between B' and C' .

The area $B'A'$ will therefore be layovered, as always happens at vertical building walls facing towards the sensor. This will lead to a mixture of signal contributions from the building and the ground in the SAR image, because the elevated objects will be closer to the sensor than the ground.

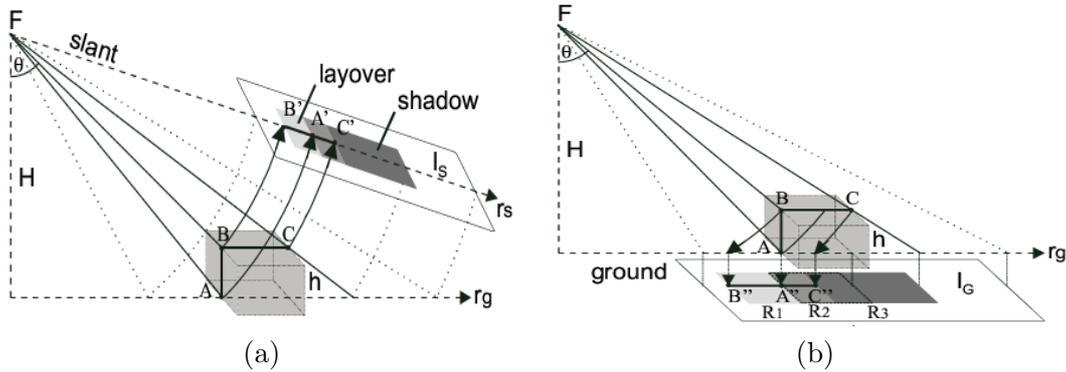


Figure 2.17: Projection of a building into a) slant and b) ground range image (Stilla et al., 2002).

Since the slant image shows a geometric distortion in the ground range direction which makes an object recognition and interpretation more difficult, for image interpretation the data will be sampled to a rectangular grid on the ground (Fig. 2.17b). However, the order of the points B'' , A'' , C'' in the ground image will still be the same as in the slant image: no complete geocoding procedure can take place, since the true heights in the scene are unknown and cannot be derived by simple unwrapping-based processes.

On its far range side, the building will cast a shadow which might occlude smaller objects behind it and that will appear as an energy-less, very low SNR, dark area in the image. In theory, the height of a detached building might be derived from the shadow length and the viewing angle. In real cases, the fact that triple reflections extend often to shadow areas poses serious limits on this kind of approach.

2.3.2 Metric SAR statistics in urban environments

When resolution cells are not anymore populated by a large number of scatterers but instead only a small number of scatterers determines their behavior, the fully developed speckle assumption (Tur et al., 1982) does not hold any more. New distributions have to be taken into account if a proper description of the statistics of the data is sought.

Usually, a K law is used to describe high resolution SAR data. The K law is derived under the hypotheses of Gamma reflectivity and Gamma speckle:

$$p_K(A = a) = \frac{2\lambda L}{\Gamma(\alpha)\Gamma(L)} (\lambda La)^{(\alpha+L)/2-1} K_{\alpha-L}(2\sqrt{\lambda La})$$

where K_n denotes the modified Bessel function of the third kind and order n and x is the amplitude. The difficulty in using K -law distributed random variables lies in the complications related to the evaluation of the Bessel function.

Inverse Gaussian distributed backscatter values corrupted by square-root Gamma speckle model well the amplitude return from the slowly varying small number of scatterers that compose these elements. They have been shown to approximate well a K -distribution

while keeping a better tractability, and have been used to model backscatter in urban environments (Frery et al., 1997; Mejail et al., 2001):

$$p(A = a) = \exp\left[-\frac{L^L \Gamma(L - \alpha) a^{L-1}}{\gamma^\alpha \Gamma(-\alpha)(\gamma + L a)^{L-\alpha}}\right] \quad (2.4)$$

with $L > 0$ the number of looks of the data, $a > 0$ the amplitude pixel value, and $-\alpha_s, \gamma_s > 0$ and $-\alpha_0, \gamma_0 > 0$ respectively the parameters for the distribution that models the amplitude values of the strong scattering elements, and those for remaining clutter obtained by solving the system

$$\begin{aligned} \gamma &= L \left[\frac{\hat{m}_1 \Gamma(-\alpha) \Gamma(L)}{\Gamma(-\alpha - 1/2) \Gamma(L + 1/2)} \right]^2 \\ \frac{\hat{m}_{1/2}^2}{\hat{m}_1} &= \frac{\Gamma^2(-\alpha - 1/4) \Gamma^2(L + 1/4)}{\Gamma(-\alpha - 1/2) \Gamma(L + 1/2) \Gamma(-\alpha) \Gamma(L)} \end{aligned}$$

where \hat{m}_n is the n -th order moment estimated on a training area and $\Gamma(\cdot)$ is the Gamma function.

2.3.3 High resolution urban SAR and InSAR inversion: state of the art

The problem of scene understanding from meter resolution SAR data is increasingly present in the remote sensing literature. Most studies limit themselves to the two-dimensional intensity case. Up to now the problem of the reconstruction in the three dimensions of urban scenes from interferometric radar data has never been addressed in a unified, consistent framework by well-integrated and clearly justified algorithms and approaches. Actually, if the peculiarities of urban HR SAR are analyzed, it is clear that the usual approach of identifying and excluding from further processing problematic shadowing or layover areas proves problematic, since they are almost ubiquitous.

Simulation based scene understanding A number of simulation-based approaches have been considered. In them, a model of the scene is progressively updated looking for the configuration that best explains the data.

The comparison between the scene and the data is actually carried out in the data space: the candidate scene is mapped into the data by using a simulator, that is able to mimic the acquisition system. The distance between data and simulated scene is computed as a cross-correlation.

In (Bolter, 2000) and (Bolter and Leberl, 2000) interferometric coherence, absolute phase and shadowing are identified as powerful image features for the reconstruction of buildings. A deterministic correlation measure between real and simulated data is used as a quality criterion for the reconstruction.

In (Gamba and Houshmand, 1999) segmentation and plane-fitting techniques are applied to interferometry-derived local elevation data to identify buildings. In one case only (Hurtas et al., 1998), a segmentation of local heights from interferometric radar is used as a cue in a pre-processing step of the building reconstruction system chain.

Ad hoc data regularization techniques A number of studies were carried out on the peculiar effects encountered in very high resolution SAR images in urban environments. They have the great merit of having brought to a better understanding of the characteristics of HR SAR phenomenology.

In (Burkhart et al., 1996) layover in the height profiles of buildings reconstructed by SAR interferometry is investigated. Some ad-hoc algorithms for the correction of this effect are also proposed. The peculiarities of HR urban SAR tend to be treated as shortcomings of the data that must be regularized instead of being used as additional sources of information.

Stochastic approaches The main limit of deterministic simulation based approaches is that

- only the scene that best explains a given dataset is extracted from the data: no full posterior probabilistic distribution for it is generated;
- they are strongly dependent on the quality of the simulator: if good results are to be obtained, a good simulator is needed, hence
- they are computationally expensive.

As an alternative, a probabilistic approach can be considered, where full probabilistic descriptions are given for the phenomena and terms involved.

To deal with the high complexity of the problem, stochastic algorithms have been proposed in recent years from a number of authors. Although exploiting interferometric radar data in a limited way, (Heuel and Nevatia, 1995) makes use in an explicit way of models for the scene objects of interest. Parallelepipedal buildings are extracted from high resolution optical images supplementing in some cases automatic processing with user intervention (Li et al., 1999; Chellappa et al., 1993; Heuel and Nevatia, 1995).

If a Bayesian system is considered, then A-Priori descriptions of the scene can be incorporated into the description.

In (Simonetto, 2002) the phenomenology of HR SAR is investigated. A stochastic algorithm for the exploitation of stereo intensity observations for industrial building extraction is proposed. In (Tupin and Roux, 2003), the extraction of building outlines using a pair of optical and SAR images is considered. In (Stoica et al., 2000), a model for a set of segments is employed in a stochastic geometry framework to overcome the difficulties with a part of the linear feature extraction process. In a related paper (Ortner et al., 2002) a similar approach is used for the reconstruction of building areas from Laser digital elevation data. In (Stilla et al., 2003), techniques for the extraction of building models from interferometric SAR data are considered and the results are compared with those obtained by LIDAR data. The full hierarchical modeling of both the geometry and the radiometry of HR SAR data acquired on complex structures in urban environments has — to the best of our knowledge — never been addressed as a tool for scene understanding, though.

2.4 Summary

In this Chapter, the following points have been discussed:

- SAR and InSAR systems have very peculiar geometric and radiometric properties that can be treated by appropriate, well established processing techniques;
- their performance at metric resolutions in urban environments is determined by
 - geometric effects such as layover, shadowing and occlusion;
 - the radiometric effect of smooth surfaces on speckle noise, backscatter texture and signal-to-noise ratio;
 - the appearance of multiple reflections of the incident radar beam;
 - the various classes of strong isolated scatterers;
 - the peculiar statistics of meter resolution data;
- the limiting effects on the traditional interferometry-based scene reconstruction approaches of the presence of very strong reflectors and of phenomena like layover and multiple reflections depend on the scene geometry. Therefore, they contain information about it;
- traditional, simulation- and model-based inversion algorithms in the literature have been evaluated and compared.

The phenomenology and effects described in this chapter form the basis for the hierarchical probabilistic modelling and inversion algorithms presented in chapters from 4 to 6 and validated in chapter III.

Chapter 3

Bayesian modeling, estimation and decision theory for multidimensional signal analysis

Abstract

Bayesian modeling and estimation techniques for the analysis of multidimensional fields are introduced. After presenting the properties of Gibbs–Markov fields on random variable networks and the results of the Hammersley–Clifford theorem, hierarchical Bayesian models are introduced.

The second level of Bayesian analysis for model selection is presented together with the role of the Occam factor in model choice and analysis.

Bayesian estimation and decision theories are then introduced: the analysis focuses on Maximum A Posteriori estimates and on decisions as posterior ratios.

Posterior optimization techniques based on complete enumeration, on Iterated Conditional Modes and on the Gibbs Sampler and Monte Carlo Markov chains are then introduced and detailed.

3.1 Bayesian modeling and analysis

The high complexity of models made up by large numbers of possibly interdependent variables can be tackled by considering and exploiting conditional independence assumptions. The use of independence assumptions allows efficient inference to be carried out even for problems involving a large number of variables.

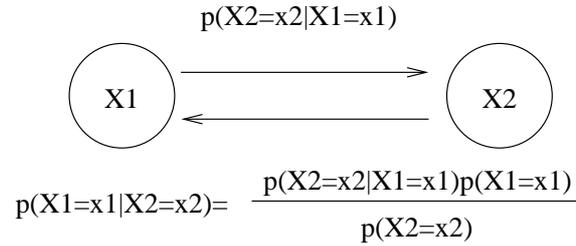


Figure 3.1: The Bayes' law enables the reversal of probabilistic links and therefore allows inference to take place based on the existence of direct models. Following the usual notation for Bayesian networks, random variables are represented by circles at the nodes of a graph while probabilistic relations are represented by (oriented) edges.

3.1.1 A principled approach to the composition of stochastic descriptions

The Bayes' equation In Bayesian probability theory, logical links are expressed by means of conditional probabilities

$$p(X_1 = x_1 | X_2 = x_2) \equiv \frac{p(X_1 = x_1, X_2 = x_2)}{p(X_2 = x_2)}$$

that express the degree of belief that event x_1 takes place given that event x_2 happens with certainty.

An immediate consequence of the definition of conditional probability is the so-called Bayes' law

$$p(X_1 = x_1 | X_2 = x_2) = \frac{p(X_2 = x_2 | X_1 = x_1) p(X_1 = x_1)}{p(X_2 = x_2)} \quad (3.1)$$

that enables the reversal of probabilistic links and therefore allows inference to take place based on the existence of direct models. The law can be seen as a rule for updating an existing description of a phenomenon X_1 , the prior $p(X_1 = x_1)$, based on new information — new data or a new description — x_2 . The direct link from the old to the new description needs to be modeled in the likelihood $p(X_2 | X_1)$. Furthermore, an evidence normalization term is introduced in $p(X_2 = x_2)$ that expresses the probability of the data and that can be computed as a Marginalization on the prior model used

$$p(X_2 = x_2) = \sum_{x_{1,i}} p(X_2 = x_2 | X_1 = x_{1,i}) p(X_1 = x_{1,i}) . \quad (3.2)$$

Bayesian inversion results in a powerful method in the understanding of complex phenomena for which precise mathematical modeling is available starting from incomplete information available in datasets of different nature.

This measure and the dependence description itself is updated from a starting prior belief to a posterior one by considering the data available.

On the logical nature of conditional probability connections It is important to notice that conditional probabilities are used to express logical links, not causal ones. If, for instance, event X_1 generates by some physical process event X_2 , a link that is both causal and logical exists between events X_1 and X_2 , but a link that is only logical exists between events X_2 and X_1 in the sense that an occurrence of event X_2 gives us a hint about the occurrence of event X_1 .

3.1.2 Modeling interdependence in random variable sets: chain Rule and Markov property

The chain rule We consider a set of N stochastic variables

$$X = (X_1, X_2, \dots, X_N)$$

with values

$$x = (x_1, x_2, \dots, x_N) .$$

Independently of the type of stochastic process considered, the joint probability of X can be written as

$$p(X = x) = p(X_1 = x_1, \dots, X_N = x_N)$$

and can be decomposed based only on the definition of conditional probability as in

$$p(X = x) = p(X_1 = x_1) p(X_2 = x_2 | X_1 = x_1) p(X_3 = x_3 | X_2 = x_2, X_1 = x_1) \\ \dots p(X_N = x_N | X_{N-1} = x_{N-1}, \dots, X_2 = x_2, X_1 = x_1)$$

which is called the “chain rule” decomposition theorem of probability. The last terms in the decomposition are based on a very large number of variables, and are therefore likely to render an inference process based on them hard to treat.

The Markov property A k -order Markov source is a symbol-emitting device with memory for which if X denotes the ordered set of symbols emitted by the source, the Markov property

$$p(X = x) = p(X_1 = x_1) \dots p(X_{k-1} = x_{k-1} | X_{k-2} = x_{k-2} \dots X_1 = x_1) \\ \prod_{i=k+1}^N p(X_i = x_i | X_{i-1} = x_{i-1} \dots X_{i-k} = x_{i-k})$$

holds. Even if the number of elements in the set X grows, only a maximum number of k symbols has to be taken into account in modeling the source.

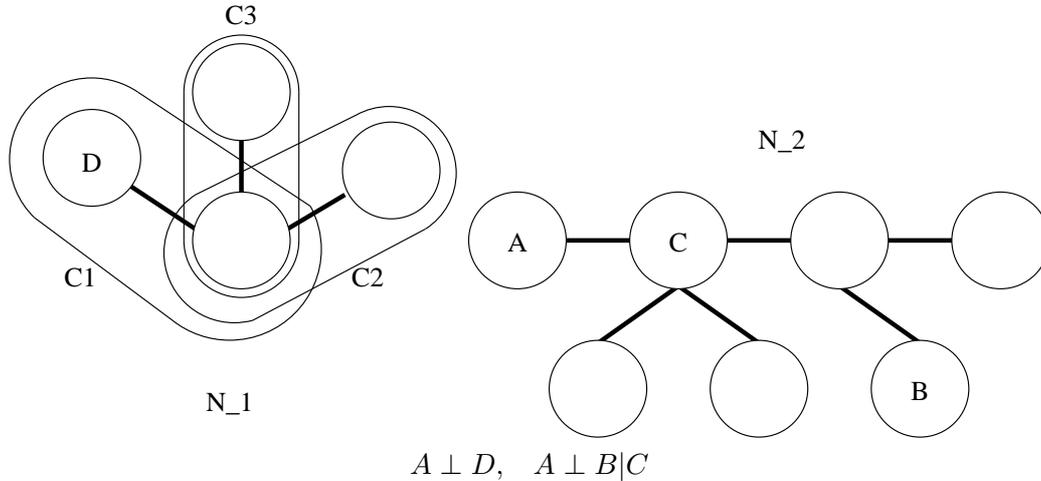


Figure 3.2: Markov property for random variables: random variables are depicted as circles. Two neighborhoods are represented. First-order cliques for neighborhood N_1 are depicted by closed shapes enclosing the variables. The three cliques C_1 , C_2 and C_3 form the first-order neighborhood N_1 . The diagram describes a belief network encoding statistical independence relations: variable A is independent of D (no path exists between the two), while A is independent of B if C is known. Independence and conditional independence are respectively denoted by $\cdot \perp \cdot$ and $\cdot \perp \cdot | \cdot$. The exploitation of conditional independence relations allows the simplification of high complexity probabilistic descriptions.

3.2 Gibbs random fields

3.2.1 Lattices, sets, neighborhoods, cliques, sites, pixel values

For any variable X_i in X a neighborhood N_i is defined by

$$N_i = \{j\} \text{ such that } \begin{cases} i \notin N_i \\ j \in N_i \Leftrightarrow i \in N_j \end{cases} .$$

A neighborhood is assigned a relative set C of cliques C_k

$$C = \{C_k\}$$

each made of k random variables and each characterized by a vector parameter

$$\vec{\theta} = \{\theta_{11}, \theta_{12}, \theta_{21}, \theta_{22}, \dots\}$$

that expresses the local interactions in the clique C_k .

While in the case of image processing and analysis neighborhoods are limited to the geometric vicinity of the central pixel under analysis X_i , it must be noticed that once again the important point is the logical link between the variables in the same clique which might or might not be associated to a location and even belong to spaces with different dimensionality.

3.2.2 Random fields in terms of potentials and energies

Local interactions, that is interactions that take place between elements of the same clique, can be now expressed in terms of energies and potentials

$$U(X = x) = \sum_{c \in C} V_c(x)$$

where the energy of the full variable set is expressed by a sum of the potentials of all the single cliques in the model and the energy of a single variable

$$U_i(X_i = x_i) = \sum_{c \in C, i \in c} V_c(x_i)$$

is expressed as a sum of all the energies of the cliques that include that variable.

Markov on a lattice: restating the Markov property at higher dimensionality

The Markov property can then be restated as specifying the interdependence between variables to actually occur only among variables belonging to the same clique. In a Markov random field a random variable can be expressed in terms of a limited number of its neighbors:

$$\text{if } x^i = (x_j)_{j \neq i} \text{ then } p(X_i = x_i | X^i = x^i) = p(X_i = x_i | x_j \in N_i).$$

The specification of statistical independence relations between random variables allows the design of faster inference algorithms taking into account only meaningful logical links.

Gibbs random fields In a Gibbs random field probability distribution functions can be expressed in terms of potentials, as in

$$p(X = x) = \frac{1}{\sum_{x_i} \exp(-U(x_i))} \exp\left(-\sum_{c \in C} V_c(x)\right).$$

The denominator partition function integral over the configuration space is often hard to evaluate due to the large number of elements in the domain of X .

The Hammersley–Clifford theorem The Hammersley–Clifford theorem states that the probability of one random variable instead of the whole set can be derived from its neighbors only, and hence states the equivalence between Markov and Gibbs fields:

$$p(X_i = x_i | X_j = x_j) = \frac{1}{\sum_{x_i, x_j} \exp(-U_i(x_i | x_j \in N_i))} \exp(-U_i(x_i | x_j \in N_i)).$$

3.2.3 Markov random fields in terms of belief networks

A Bayesian network (Hackerman et al., 1995) is a graphical model for probabilistic relationships among variables in a set. It consists of an annotated directed graph that encodes

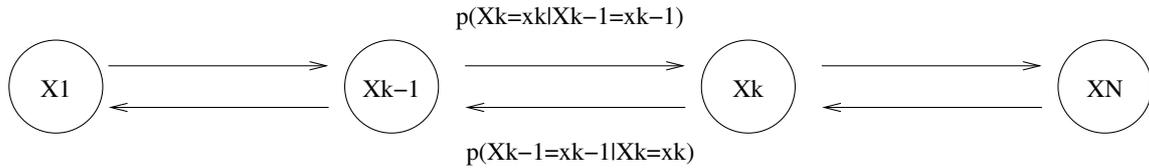


Figure 3.3: Hierarchical model illustration: the distribution of parameters at any level of the hierarchy depends on parameters at the next lower level and, conditional on those parameters, is independent of parameters at all levels below that.

probabilistic relationships among variables of interest in uncertain-reasoning problems. The representation formally encodes the joint probability distribution for its domain, yet includes a human-oriented qualitative structure that facilitates communication between a user and a system incorporating the probabilistic model.

Bayesian networks allow us to overcome the curse of dimensionality by encoding probabilistic independence between the elements of the inference problem. They can be mapped to graphs whose nodes represent the random variables under study, while dependencies are encoded in their structure. The learning procedure progressively updates the network thereby obtaining a description of the statistical link between the variables based on all the information available. Once the graph structure is set up, it can be used to perform efficient inference by taking into account only significant dependencies in between the random variables.

The nodes in the graph correspond to the variables in the problem, while the directed edges encode the significant probabilistic dependencies. The fundamental property of the net is that a node is conditionally independent of its ancestors given its parents. Hence, inference on the graph can be reduced to a message passing algorithm in which starting from a set of root nodes, probability is propagated down the graph to the node of interest.

3.2.4 Hierarchical models

As noted by O'Hagan (1994), the $p(X_1 = x_1 | X_2 = x_2)$ posterior statistical model and $p(X_2 = x_2)$ prior model together form an ordered structure in which the distribution of the data x_1 is written conditionally on parameters $X_2 = x_2$ as $p(X_1 = x_1 | X_2 = x_2)$. Analogously, the prior distribution of X_2 can be written conditionally on hyperparameters X_3 as $p(X_2 = x_2 | X_3 = x_3)$, and be completed by the distribution of $X_3 = x_3$, $p(X_3 = x_3)$. We could go further and write the distribution of X_3 conditionally on some more ('hyper-hyper') parameters $X_4 = x_4$ as $p(X_3 = x_3 | X_4 = x_4)$, and this process could continue as far as is needed. Such models are called *hierarchical* models, because of the way in which the distribution of parameters in each level of the hierarchy depends on the parameters in the previous level.

In fact it is usual to say that the distribution of parameters at any level of the hierarchy depends on parameters at the next lower level *and*, conditional on those parameters, is independent of parameters at all levels below that. For instance, if we model the distribution

of X_2 in terms of $p(X_2 = x_2|X_3 = x_3)$ and $p(X_3 = x_3)$, then the likelihood $p(X_1 = x_1|X_2 = x_2)$ is formally the distribution of X_1 given $X_2 = x_2$ and $X_3 = x_3$. Writing it as $p(X_1 = x_1|X_2 = x_2)$ incorporates a judgment that if we know $X_2 = x_2$ then knowing $X_3 = x_3$ would not add any information about X_1 . This is reasonable because X_3 has been introduced only as a way of formulating $p(X_3 = x_3)$ as in the (discrete) evidence integral

$$p(X_2 = x_2) = \sum_{x_3} p(X_2 = x_2|X_3 = x_3)p(X_3 = x_3) . \quad (3.3)$$

The reason for making this interpretation of $p(X_1 = x_1|X_2 = x_2)$ is that otherwise the distributions of $p(X_1 = x_1|X_2 = x_2)$, $p(X_2 = x_2|X_3 = x_3)$ and $p(X_3 = x_3)$ together do not completely specify the joint distribution of X_1 , X_2 and X_3 . The extra assumption allows us to write

$$p(X_1 = x_1, X_2 = x_2, X_3 = x_3) = p(X_1 = x_1|X_2 = x_2) p(X_2 = x_2|X_3 = x_3) p(X_3 = x_3) . \quad (3.4)$$

A hierarchical model does always specify the full joint distribution of all quantities this way ¹.

3.2.5 Principle of inference for 2-level hierarchical models

If a set $(\mathcal{M}_1, \dots, \mathcal{M}_N)$ of models is available, and we consider that each model $\mathcal{M} = m$ has a vector of parameters $\Theta = \theta$, then a full three-levels hierarchical model can be described as in

$$\mathcal{M} \rightarrow \Theta \rightarrow D .$$

A model is defined by its functional form and two probability distributions: a prior distribution $p(\Theta = \theta|\mathcal{M} = m)$ which states what values the model's parameters might plausibly take; and the predictions $p(D = d|\Theta = \theta; \mathcal{M} = m)$ that the model makes about the data $D = d$ when its parameters have a particular value θ . Note that models with the same parameterization but different priors over the parameters are therefore defined to be different models.

At the first level of inference, we assume that one model $\mathcal{M} = m$ is true, and we infer what the model's parameters $\Theta = \theta$ might be given the data d . Using Bayes' rule in eq. 3.1, the posterior probability of the parameters θ is:

$$p(\Theta = \theta|D = d; \mathcal{M} = m) = \frac{p(D = d|\Theta = \theta; \mathcal{M} = m) p(\Theta = \theta|\mathcal{M} = m)}{p(D = d|\mathcal{M} = m)} . \quad (3.5)$$

¹The equivalence of this formulation with that of Bayesian networks and graphical models is indicated by the correspondence of the description above with the textbook definition (as in Heckerman (1995) or in Myllymaki and Tirri (1993)) of a Bayesian belief network representation for a probability distribution $p(\cdot)$ on a domain (X_1, \dots, X_N) as a pair (G, P) where G is a directed acyclic graph whose nodes correspond to the variables X_1, \dots, X_N and whose topology satisfies the following: each variable X_i is conditionally independent of all of its non-descendants in G , given its set of parents pa_{X_i} and no proper subset of pa_{X_i} satisfies this condition. The second component P is a set consisting of all the conditional probabilities of the form $p(X_i|pa_{X_i})$.

The normalizing constant $p(D = d|\mathcal{M} = m)$ is commonly ignored, since it is irrelevant to the first level of inference, i.e., the choice of $\Theta = \theta$; but it is important in the second level of inference, and is called the evidence. It is common to use gradient-based methods to find the maximum of the posterior, which defines the most probable value for the vector parameter, $\hat{\theta}_{\text{MAP}}$, that is subsequently used to summarize the posterior distribution by its value ² (section 3.3).

3.2.6 Occam razor and Occam factor

As noted by MacKay (1992), model comparison is a difficult task because it is not possible simply to choose the model that fits the data best: more complex models can always fit the data better, so the maximum likelihood model choice would lead us inevitably to implausibly over-parameterized models which generalize poorly.

‘Occam’s razor’ is the principle that states that unnecessarily complex models should not be preferred to simpler ones. Bayesian methods automatically and quantitatively embody Occam’s razor (Jeffreys, 1939; Gull, 1988), without the introduction of ad hoc penalty terms: complex models are automatically self-penalizing under Bayes’ rule.

At the second level of inference, we wish to infer which model is most plausible given the data. The posterior probability of each model is:

$$p(\mathcal{M} = m|D = d) \propto p(D = d|\mathcal{M} = m) p(\mathcal{M} = m) . \quad (3.7)$$

Notice that the data-dependent likelihood term $p(D = d|\mathcal{M} = m)$ is the first-level Bayesian inference evidence for $\mathcal{M} = m$ of equation 3.5. The second term, $p(\mathcal{M} = m)$, is a ‘subjective’ prior over the hypothesis space, kept constant when there is no reason to assign strongly differing priors $p(\mathcal{M} = m)$ to the alternative models.

The evidence then naturally embodies Occam’s razor: in order to assign a preference to alternative models $\mathcal{M} = m$, the evidence $p(D = d|\mathcal{M} = m)$ has to be evaluated.

It should be pointed out that the emphasis of this modern Bayesian approach is not on the inclusion of priors into inference. Historically, Bayesian analysis has been accompanied by methods to work out the ‘right’ prior $p(\Theta = \theta|\mathcal{M} = m)$ for a problem, for example, the principles of insufficient reason and maximum entropy. Modern Bayesian theory, however, does not take a fundamentalist attitude to assigning the ‘right’ priors – many different priors can be tried, allowing the data to inform us which is most appropriate. Each particular prior corresponds to a different hypothesis about the way the world is. These alternative hypotheses can be compared in the light of the data by evaluating the evidence.

²The width of the posterior distribution is approximately characterized by the curvature of the posterior; writing the Hessian $\mathbf{H} = \nabla\nabla \log p(\Theta = \theta|D = d; \mathcal{M} = m)$ and Taylor-expanding the log posterior with $\Delta\theta = \theta - \hat{\theta}_{\text{MAP}}$, with T denoting the transpose operator,

$$p(\Theta = \theta|D = d; \mathcal{M} = m) \simeq p(\hat{\theta}_{\text{MAP}}|D = d; \mathcal{M} = m) \exp\left[-\frac{1}{2} \Delta\theta^T \mathbf{H} \Delta\theta\right] \quad (3.6)$$

we see that the posterior can be locally approximated as a Gaussian with covariance matrix \mathbf{H}^{-1} .

In practice, the evaluation of the Marginalization integral of eq. 3.3

$$p(D = d|\mathcal{M} = m) = \sum_{\theta} p(D = d|\Theta = \theta, \mathcal{M} = m) p(\Theta = \theta|\mathcal{M} = m) \quad (3.8)$$

can be conducted by approximating the posterior as a Gaussian around its MAP peak, using Laplace's method:

$$p(D = d|\mathcal{M} = m) \simeq p(D = d|\hat{\theta}_{\text{MAP}}, \mathcal{M} = m) \cdot p(\hat{\theta}_{\text{MAP}}|\mathcal{M} = m) \det^{-\frac{1}{2}}(\mathbf{H}/2\pi) \quad (3.9)$$

in which the Hessian $\mathbf{H} = -\nabla\nabla \log p(\Theta = \theta|D = d; \mathcal{M} = m)$ appears in the last two terms that account for the so called 'Occam factor'

$$\Omega \equiv p(\hat{\theta}_{\text{MAP}}|\mathcal{M} = m) \det^{-\frac{1}{2}}(\mathbf{H}/2\pi) < 1 \quad (3.10)$$

which penalizes $\mathcal{M} = m$ for having the parameter θ . The first-level evidence $p(D = d|\mathcal{M} = m)$ can then be obtained simply by multiplying the best fit likelihood by the Occam factor, favoring less complicated models whenever a choice is made for instance by making use of the Bayes' factor

$$B_{ij} = \frac{p(D = d|\mathcal{M} = m)}{p(D = d|\mathcal{M} = m')} \simeq \frac{\Omega_m}{\Omega_{m'}} \quad (3.11)$$

in a set of given models.

The maximization of the evidence is therefore regarded as a criterion for the choice of a suitable model to explain a given dataset.

Jeffrey's priors Frequently partial prior information is available, outside of which it is desired to use a prior that is as non-informative as possible.

The maximum entropy distribution "is uniquely determined as the one which is maximally noncommittal with regard to missing information" (Jaynes, 1957): it "agrees with what is known, but expresses 'maximum uncertainty' with regard to all other matters, and thus leaves a maximum possible freedom for our final decision to be influenced by the subsequent sample data" (Jaynes, 1968). It has been further shown by Jaynes that the maximum entropy distribution is equal to the frequency distribution that can be realized in the greatest number of ways.

In order to define the requested maximum entropy prior, constraints can be composed by using Lagrange multipliers with a Shannon entropy measure

$$-\sum_x p(X = x) \log p(X = x) \quad (3.12)$$

and the resulting composite equation can therefore be maximized providing the requested resulting prior.

3.2.7 From posterior distributions to estimates and decisions

Full probabilistic inference produces as output full posterior distributions rather than simple scalar values. On the other hand, in a number of situations, scalar values are required in order to be able to compare a theoretical result with actual measured quantities.

Estimation and decision theory considers the problem of choosing a representative value for a given probabilistic distribution in a principled way.

3.3 Estimation and decision theory

The problem of estimation and decision theory is the derivation in a principled way of criteria for the choice of single scalar values from a full probabilistic description of a problem.

It is clear that moving from a full probabilistic description of a phenomenon to a single value must involve some kind of loss in the quality of the descriptor. This loss must be expressed by some kind of cost function leading to a definition of an optimal estimator: the classical parameter estimation is formulated as the minimization of the Bayes risk R defined over the signal space. The risk R is the expectation value of a cost function defined over the joint space of observations (Y) and model parameters X

$$R = E[c(X = x, \hat{x})] = \sum_y \sum_y c(X = x, \hat{x}) p(X = x, Y = y) . \quad (3.13)$$

The cost function $c(X = x, \hat{x})$ is a measure of goodness of the estimated parameter \hat{x} , being defined as a distance between the actual and the desired but unknown value of x . Minimizing the risk and using a uniform cost function, at the limit $\epsilon \rightarrow 0$, we obtain the Maximum A Posteriori (MAP) estimator that summarizes the posterior description of phenomenon y in terms of the position of its maximum

$$\hat{y}_{\text{MAP}} = \arg \max_y \{p(X = x|Y = y) p(Y = y)/p(X = x)\} . \quad (3.14)$$

We observe that, in classical estimation theory, using a cost function is nothing else but describing a type of prior information. The expression for the posterior encapsulates the deterministic prior knowledge represented by the forward model. In addition, the knowledge about the observation noise and the a priori information about the desired parameter are also included. We conclude that MAP is a complete frame for model-based approaches in information extraction.

This can be demonstrated (Rissanen, 1985) to be equivalent to a shortest description length estimate obtained by considering that the best model of a phenomenon is the one that produces the most compact encoding of it. A very similar approach, also considering two terms, a data one requiring the maximization of a likelihood and a penalty term considering the complexity of the model, is the Akaike information criterion (Akaike, 1974).

3.3.1 Decisions as posterior ratios

Decision theory focuses on the problem of choosing the most probable hypothesis in a set of given ones. It is clearly strongly related to the problem of model selection as well as to estimation: even in this case, the fundamental problem is the reduction of a full probabilistic description to a single value. We consider the problem of binary hypothesis testing under data D . If the two competing hypotheses we are considering are called H_0 and H_1 , it can be seen that the decision reduces itself to the evaluation of the ratio

$$\frac{p(H = h_0)p(D = d|H = h_0)}{p(H = h_1)p(D = d|H = h_1)}$$

that, introducing the evidence $p(D)$ and using the Bayes' theorem, equals

$$\frac{p(H = h_0|D = d)}{p(H = h_1|D = d)}$$

a comparison of the posterior probabilities that becomes, when identical priors are considered, to the usual likelihood ratio estimator of decision theory

$$L(D = d) \equiv \frac{p(H = h_0|D = d)}{p(H = h_1|D = d)}.$$

3.3.2 Estimation and decision algorithms

The optimization of probability distributions of large sets of random variables is complicated by the large dimensionality of the spaces where these sets are located. Examples of such sets include

- gray-scale image processing: X is a set of mono-dimensional stochastic pixels X_i with values x_i in some set L . The number of configurations N is usually very large (it easily takes up values in the order of a million), and the size of L can be considerably large as well, especially if instead of limiting the problem to a quantized space of 256 values larger pixel ranges are considered
- color image processing: the size of L increases considerably with respect to the gray-scale image processing example. Quite often, different mechanisms have to be accounted for in describing the way different physical phenomena are responsible for the generation of the different bands
- hierarchical geometrical modeling: X is decomposed into a series of subsets $X = (X_{1D}, X_{2D}, X_{3D})$ each of which models a set of interdependent objects with different dimensionality. For instance, the full problem space might be composed of a set of points that probabilistically generate a set of segments that are composed into bi- and three-dimensional graphs.

The characterization of the behavior of the set X is in cases like these made complex by the very size of the state space one has to consider. Even in the simple case of a 1000×1000 pixels 256-levels image processing, with N fixed and X_i mono-dimensional, the possible number of states of X is $1000 \times 1000 \times 256$.

3.3.2.1 Analytic derivation of MAP estimate

In the cases when the posterior is analytic, the MAP estimate

$$\hat{x} = \arg \max_x p(X = x|Y = y)$$

can be computed directly from

$$\frac{dp(X = x|Y = y)}{dx} \Big|_{X=\hat{x}_{\text{MAP}}} = 0, \quad \frac{d^2p(X = x|Y = y)}{dx^2} \Big|_{X=\hat{x}_{\text{MAP}}} < 0.$$

When this is not possible, the optimization is instead carried out by numerical methods (Cohen and Cooper, 1987).

3.3.2.2 Exact inference by complete enumeration

A possible method for statistical inference is a brute-force one: complete enumeration of all the hypotheses, and evaluation of their probabilities. This approach is an exact method, and the difficulty of carrying it out motivates the approximate methods introduced in later Paragraphs.

It has to be noted that, even though the hypothesis spaces of many problems can be thought of as continuous, such continuous spaces can be discretized and so can, in principle, be enumerated –at a grid of parameter values, for example.

3.3.2.3 The Iterated Conditional Modes algorithm

The Iterated Conditional Modes (Besag, 1986; Winkler, 1995) algorithm is an iterative optimization technique specifically designed for probabilistic models. It is a local technique, and so is susceptible to local minima. The algorithm works by iteratively computing the MAP estimate of the data using the available observations and the current estimates of the model parameters until convergence is reached. The IEM algorithm is closely related to the Expectation–Maximization algorithm of Dempster et al. (1977), that operates instead by optimizing on a simpler, convex approximation (a form of the Jensen’s inequality, the fact that an arithmetic mean is never smaller than a geometric mean, is used as a bound) of the posterior.

3.3.2.4 The Gibbs sampler and Monte Carlo Markov chains

A Metropolis algorithm based on a Simulated Annealing procedure (see Appendix B) can be adopted for the estimation of the MAP: the multivariate posterior is maximized by iteratively evaluating many possible solutions in a multidimensional parameter space.

The Metropolis algorithm can therefore be understood as a special case of a threshold random search, a relaxation of the greedy maximal ascent algorithm.

<ul style="list-style-type: none"> • Step 0: Choose arbitrary starting point $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_p^0)'$ and set $i = 0$; • Step 1: Generate $\theta^1 = (\theta_1^1, \theta_2^1, \dots, \theta_p^1)'$ as follows: <ul style="list-style-type: none"> – Generate $\theta_1^{i+1} \sim p(\Theta_1 = \theta_1 \Theta_2 = \theta_2^i, \Theta_3 = \theta_3^i, \dots, \Theta_p = \theta_p^i, D)$; – Generate $\theta_2^{i+1} \sim p(\Theta_2 = \theta_2 \Theta_1 = \theta_1^i, \Theta_3 = \theta_3^i, \dots, \Theta_p = \theta_p^i, D)$; – ... – Generate $\theta_p^{i+1} \sim p(\Theta_p = \theta_p \Theta_1 = \theta_1^i, \Theta_2 = \theta_2^i, \dots, \Theta_{p-1} = \theta_{p-1}^i, D)$; • Step 2: Set $i = i + 1$, go to Step 1.

Table 3.1: The Gibbs Sampling algorithm. Starting from a purely random initial configuration θ^0 , the algorithm sequentially updates all parameters at each iteration by sampling the new value from the local probability distribution conditioned on those that influence the one under analysis.

This approach allows the system to explore solutions that a direct maximum-gradient maximization would not consider, permitting it to identify the global maximum of a multivariate distribution avoiding the burden of constructing the full posterior distribution.

A Gibbs sampler can be set up to sample the posterior distribution keeping all the dimensions of the Scene process fixed but one: this modified distribution can be shown (Winkler, 1995; Stoica et al., 2000) to converge to the full posterior.

For generating the new candidate solutions to the optimization problem from the previous ones, update rules have to be defined. Special care must be used in ensuring that the mutation dynamics of the Gibbs sampler are completely reversible, to avoid getting stuck on some local maxima or on the borders of the domain. Each of the defined mutation step must have an equiprobable opposite move.

3.4 Summary

In this chapter, the following points have been discussed:

- Bayesian modeling and estimation techniques for the analysis of multidimensional fields have been introduced;
- the properties of Gibbs–Markov fields and hierarchical Bayesian models were detailed;
- Bayesian model selection was presented together with the role of evidence maximization and of the Occam factor in model choice and analysis;

-
- Bayesian estimation and decision theories were introduced focusing on Maximum A Posteriori estimates and on decisions as posterior ratios;
 - Posterior optimization techniques based on complete enumeration, on Iterated Conditional Modes and on the Gibbs Sampler and Monte Carlo Markov chains were then introduced and detailed.

The techniques presented in this chapter are applied to the construction of hierarchical models of complex acquisition procedures and of their results and to the derivation of estimation algorithms for the inversion of these procedures in the next chapters of this thesis.

Part II

Hierarchical Bayesian modelling

Abstract

A set of novel algorithms for urban scene understanding from high resolution SAR imagery are introduced that are based on hierarchical stochastic models of the scene and the data and on Bayesian estimation and decision.

A new algorithm for space-variant model order selection in Bayesian model-based image denoising and information extraction is developed based on evidence maximization in the framework of second-level Bayesian inference for model selection. The algorithm is demonstrated on real and simulated data generated from different models and corrupted by multiplicative speckle noise.

Based on the information extracted and in order to complement different types of information derived from various data sources, a novel approach to scene reconstruction based on Bayesian data classification and fusion is introduced: generic features extracted from the data are combined with user conjectures to generate land use maps. The system is demonstrated on real InSAR data acquired by spaceborne sensors.

To overcome the limitations of pixel-based analysis applied to sub-metric resolution data, an new object-based algorithm is presented that models the scene as an instance of a decomposable-object hierarchic marked-point prior process. The geometric and radiometric phenomenologies of sub-metric SAR in urban environments are taken into account in the likelihood term. The optimization of the obtained posterior is based on Gibbs sampling and Monte Carlo Markov chains. The developed algorithm is demonstrated on real data acquired by airborne X band sensors on built-up areas in high complexity settings characterized by tall vegetation and strong metallic scatterers.

Chapter 4

Space–variant model selection in model–based information extraction

Abstract

While the problematic of model selection in the Bayesian framework is well known, most applications in the image processing literature deal with clean data that are assumed completely known.

Here instead, an extension of the image denoising and information extraction method of Walessa and Datcu (2000) is introduced that incorporates a further estimation level to allow the space–variant selection of the model order by means of the maximization of the Bayesian model evidence.

This allows the system to properly model and reconstruct images that are composed of instances of different stochastic image models with varying degrees of complexity.

The extended algorithm is demonstrated on simulated data generated with well–known model orders as well as on Brodatz textures corrupted by multiplicative speckle noise. Classification maps based on model order are generated that for airborne X band SAR data at metric resolution demonstrate the discrimination potential of the Bayesian evidence maximization with respect to model choice.

The first approach to data analysis we consider fills the gap from data to application generic signal features. In particular, we consider the problem of image denoising and information extraction. A model–based approach (Walessa, 2000) can be applied that models the image data as a sample of a Markov random field whose parameters express the local texture (Cross and Jain, 1983).

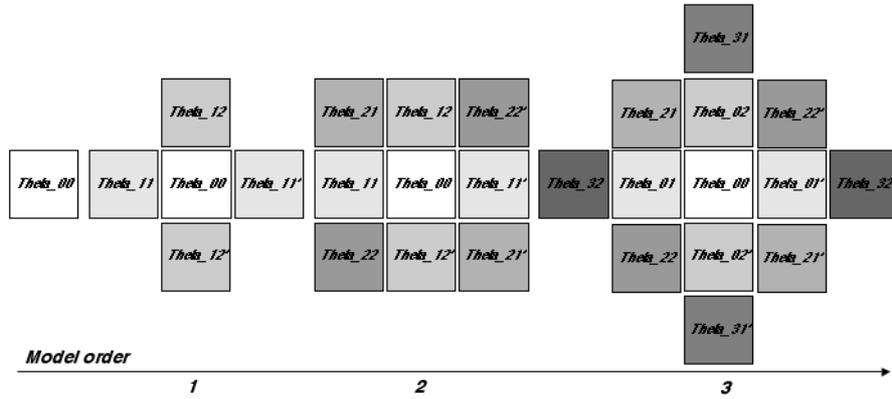


Figure 4.1: Clique structures for different Gauss–Markov random field model orders (respectively one, two, three and four): the number of model parameters in the vector θ increases with model order, allowing the expression of more complex structures.

4.1 Model based image denoising and information extraction

Bayesian model based image denoising makes use of the Bayes’ equation

$$p(X = x|Y = y) = p(Y = y|X = x) p(X = x) / p(Y = y)$$

applying it to the case where X is a stochastic process generating images x , Y describes “dirty” images y and the likelihood $p(Y = y|X = x)$ describes the noise in the image. If a further modeling level Θ generating X is included in the analysis and its most probable parameter $\hat{\theta}_{\text{MAP}}$ is estimated, the denoising system is able to perform an information extraction task: the clean image is not only extracted from the noisy one, but it is also explained in terms of a lower–level model. In this chapter, we will introduce a further extension to the speckle filtering and image information extraction approach described by Walesa and Datcu (2000), introducing a third modeling level that, in addition to the denoising and model parameter estimation, allows the local space–variant estimation of the most probable model to have generated the data.

4.2 Image models

The probabilistic models used play a central role in Bayesian analysis. We introduce here the image models and the noise models used in the denoising and information extraction system of Walesa and Datcu (2000) and Datcu et al. (1998) that is object of the extension to space-variant model order selection in the following section of this chapter.

Since the system is developed in a Bayesian framework, the choice of appropriate prior models for the clean data plays an important role, and affects in a direct way the obtained results.

A prior scene model used for scene reconstruction should ideally be completely independent of the data available and its characteristics. On the other hand, the development of tractable scale-independent probabilistic scene models is a complicated task, both for the practical need of constraining the model complexity and because the properties of the sensor (and in particular its incidence angle and spatial resolution) we influence in a very direct way the classes of objects that are actually observable and their perceived level of detail.

4.2.1 Gauss–Markov random fields

The prior model considered for image data is the Gauss-Markov random field (GMRF) (Datcu et al., 1998; Schröder et al., 1998a)

$$p(X_s = x_s | X_r = x_r, r \in G_{\mathcal{M}_s}, \sigma_s, \theta_s, \mathcal{M}_s) = \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp \left[-\frac{(x_s - \sum_{r \in G} \theta_{s,r} x_{s+r})^2}{2\sigma_s^2} \right]. \quad (4.1)$$

It expresses the generic pixel value x_s as normally distributed with width σ_s around the mean of its neighboring pixels x_{s+r} symmetrically weighted by the texture parameter vector $\theta = (\theta_0, \dots, \theta_m)$ defined on a neighborhood of cliques G centered on the pixel x_s , and such that the scalar parameters are symmetric around the central element (see figure 4.1) according to Chellappa (1985); Chellappa et al. (1985).

The main strength of the Gauss-Markov model lies in its ability to model local dependencies between pixels in a wide set of textured images, while still allowing analytical tractability.

While the prior model employed in both information extraction modules is the GMRF, the likelihoods considered have to be adapted to the characteristics of the data under analysis.

4.2.2 Gaussian likelihoods for additive Gaussian noise

If an additive Gaussian noise model is assumed for the data (as is the case, for example, for SAR interferometry-derived Digital Elevation models), a Gaussian likelihood can be used

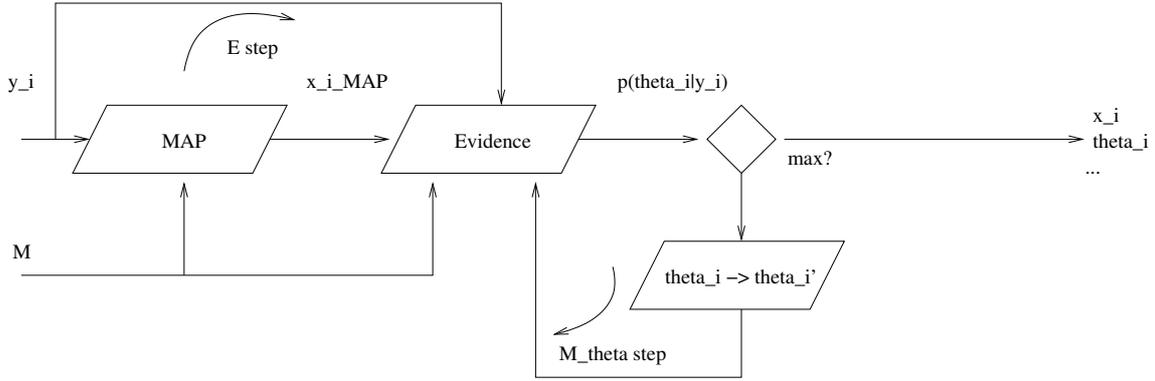


Figure 4.2: Flowchart of the evidence maximization algorithm used for texture parameter estimation in the model based denoising and information extraction system seen as an Expectation–Maximization algorithm.

to describe the relation between noise-less pixel values and their noise corrupted instances:

$$p_{\hat{h}}(Y_i = y_i | X_i = x_i) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left[-\frac{(y_i - x_i)^2}{2\sigma_n^2}\right]$$

where σ_n is the width of the distribution.

4.2.3 Gamma likelihoods for multiplicative fully developed speckle noise

The data model can be also adapted to model uncorrelated multiplicative noise, as for the (fully developed) speckle corrupting coherent SAR backscatter intensity data.

The likelihood therefore employed in the Bayes equation is a space-variant square root Gamma distribution

$$p(Y_i = y_i | X_i = x_i) = 2\left(\frac{y_i}{x_i}\right)^{2L-1} \frac{L^L}{x_i \Gamma(L)} \exp\left[-L\left(\frac{y_i}{x_i}\right)^2\right] \quad (4.2)$$

where y_i and x_i are the corrupted and the original backscatter square-root intensity values, L is the number of looks of the data, and $\Gamma(\cdot)$ is the Gamma function.

4.3 Locally-adaptive parameter estimation by evidence maximization

The denoising system of Walesa and Datcu (2000) uses the GMRF and the square-root Gamma models to describe and filter out additive Gaussian and speckle noise from images while preserving basic textures performing an inversion of the hierarchical model

$$\Theta \rightarrow X \rightarrow Y$$

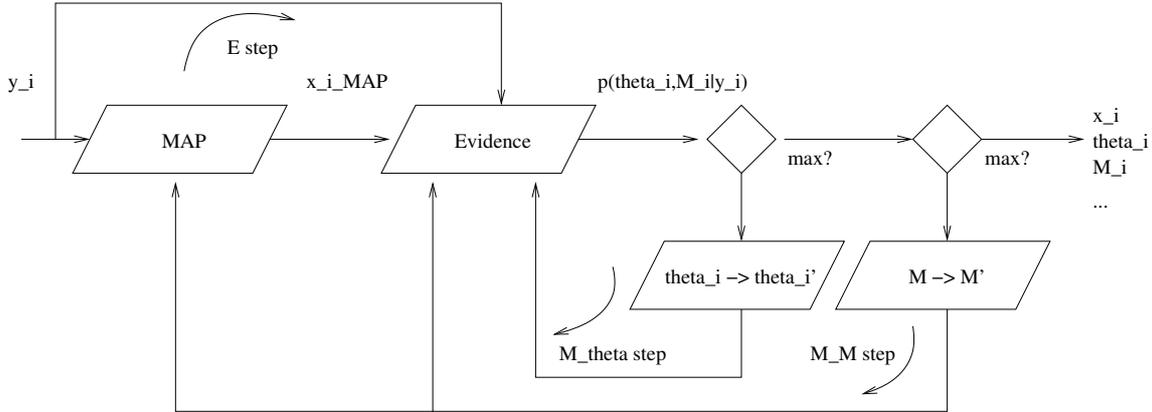


Figure 4.3: Flowchart of the evidence maximization algorithm used for contextual texture parameter and model order estimation in the extended model based denoising and information extraction system seen as an Expectation–Maximization algorithm.

being Θ , X and Y stochastic processes describing respectively clean image parameters, clean images and noise–degraded images.

Analytically computed MAP estimates of the noise-less data are generated from the filter. These estimates are then employed (as in figure 4.2) to produce parameters for the Gauss–Markov texture random field by maximization of the analytical evidence term $p(Y_i = y_i | \Theta_i = \theta_i, M = m) \propto p(\Theta_i = \theta_i | Y_i = y_i, M = m)$. The model order $\mathcal{M} = m$ is considered fixed. The estimated model parameters express the characteristics of the texture and the strength of geometrical structures in the data and can therefore be used as an input to further image interpretation methods.

Non-linear features such as sharp edges and targets are also extracted from the data and separately handled in order to restore them in the filtered images.

4.3.1 Fixed model order computational overview

Usually, stochastic relaxation methods are required for MAP optimization tasks. In the system, however, a deterministic Iterated Conditional Modes algorithm (Besag, 1986; Winkler, 1995) is employed, since convergence to the final solution is much faster. The algorithm is applied to the estimation by

- computing the MAP estimate of the despeckled backscatter using the available observations and the current estimates of the model parameters;
- keeping the MAP estimate fixed and determining the model parameters that maximize the expression for the evidence.

To be able to perform the integration over the posterior product $p(Y = y | X = x)p(X = x | \Theta = \theta)$ and to keep the problem tractable several approximations must be made (Walessa and Datcu, 2000):

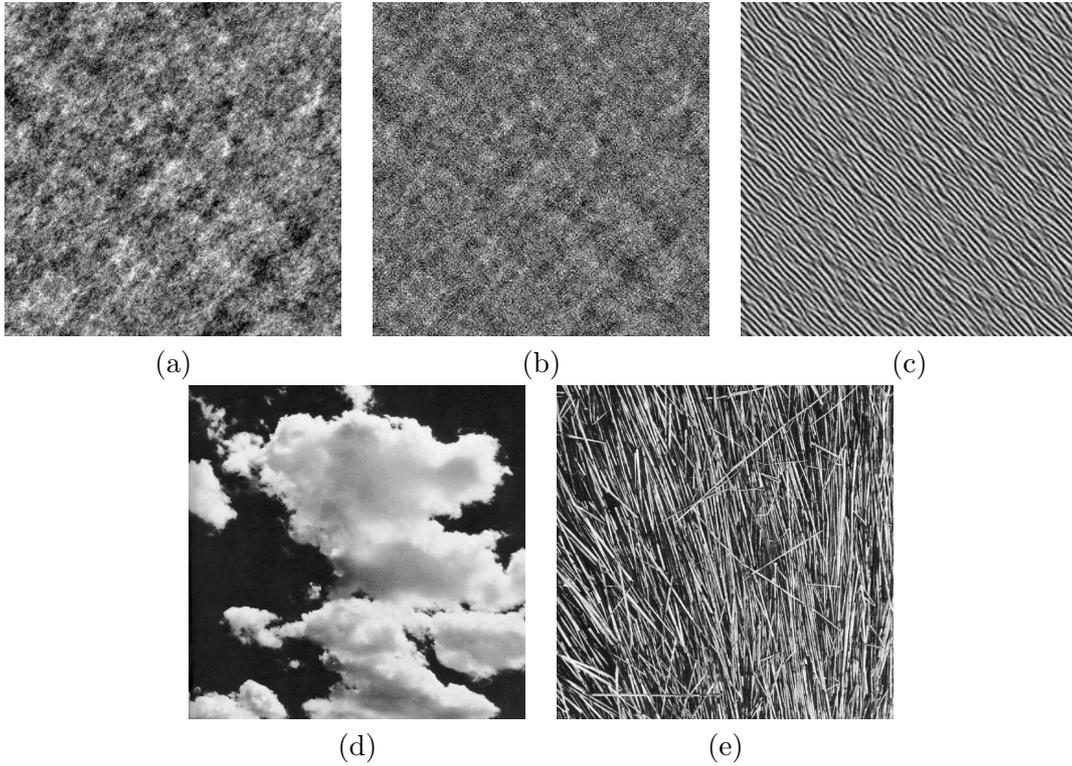


Figure 4.4: Synthetic and natural realizations of random fields. Figures (a) and (b) are instances of Gauss–Markov random fields obtained by Fourier techniques (Fenton, 1990) with model orders respectively 2 and 5. Figure (c) is an instance of a synthesized autobinomial model. Figures (d) and (e) are examples of natural textures from the Brodatz (1966) album respectively with low and high textural complexity, and therefore presumably describable respectively through low and high model order fields.

- the multi-dimensional posterior product is approximated by a multivariate Gaussian distribution, which is centered around the MAP estimate \hat{x}_{MAP} , i.e. around the maximum of the posterior (MacKay, 1992)
- the integrand of equation 3.2 is considered to consist of mutually independent random variables, allowing us to break the conditional probability density functions into the products of their components. Of course, this statistical independence is not given but has been shown to be a good approximation for large numbers of pixels, being equivalent to the maximum pseudo-likelihood approach (Besag, 1986).

By approximating the square-root Gamma likelihood with a Gaussian around its maximum, an analytic expression for the Occam factor can be computed.

4.3.2 Fixed model order processing examples and global model evidence computation

A set of test images is reported in figure 4.4. It contains synthetic and natural realizations of random fields. In particular, figures 4.4a and 4.4b are instances of Gauss–Markov random fields obtained by Fourier techniques (Fenton, 1990) with model orders respectively 2 and 5. Figure 4.4c is an instance of a synthesized autobinomial model (Cross and Jain, 1983). Figures 4.4d and 4.4e are examples of natural textures from the Brodatz album (Brodatz, 1966) respectively with low and high textural complexity, and therefore presumably describable respectively through low and high model order fields.

All the images are corrupted with a three-looks speckle noise generated according to Rainey and Wessels (1988): if L denotes the number of looks of the output image, L independent instances of Rayleigh distributed noise are generated, band-passed and averaged before being multiplied with the original noise-less image.

The generated images are fed to the filtering system imposing both low and high model orders: the results show too that low model orders generate an over-smoothing of the results, while too high model orders might generate spurious results by letting noise leak into the texture parameters of highest order.

Examples of denoising and information extraction with fixed model order are given in figure 4.6: results are shown on the model order 2 synthetic noisy texture in 4.5a: for each model $\mathcal{M} = m$ used for processing, the model order, the obtained local backscatter σ_0 and texture $||\theta||$ intensities are reported together with the local evidence of the model $p(\mathcal{M} = m)$ and its histogram with “+” marking model order 2 and “×” marking model order 5. The simpler model order 2 is most evident in the data, consistently with the considered simulation parameters. Similar results, reported in figures 4.7, hold for natural Brodatz textures.

4.4 Model order selection

It is clear that the order influences the degree of structural complexity that a model is able to express: lower order models properly describe simple, uniform scenes, but might result in an over-smoothing of very structured data; on the other hand, higher order models are able to describe more complicated dependencies at the cost of growing complexity and at the risk of introducing noise-induced artifacts when dealing with simple, unstructured data.

Since both natural and remote sensing images are in general only approximated by instances of Gauss–Markov random fields, imposing a uniform model order to a whole image can severely limit the quality of the obtained results. Conversely, a system for the automatic analysis of local image model order can be used to discriminate between phenomena of different complexity in the data.

To be able to distinguish between different model orders in the data, a new modeling level needs to be added to the hierarchy leading from the GMRF parameter $\Theta = \theta$ to the

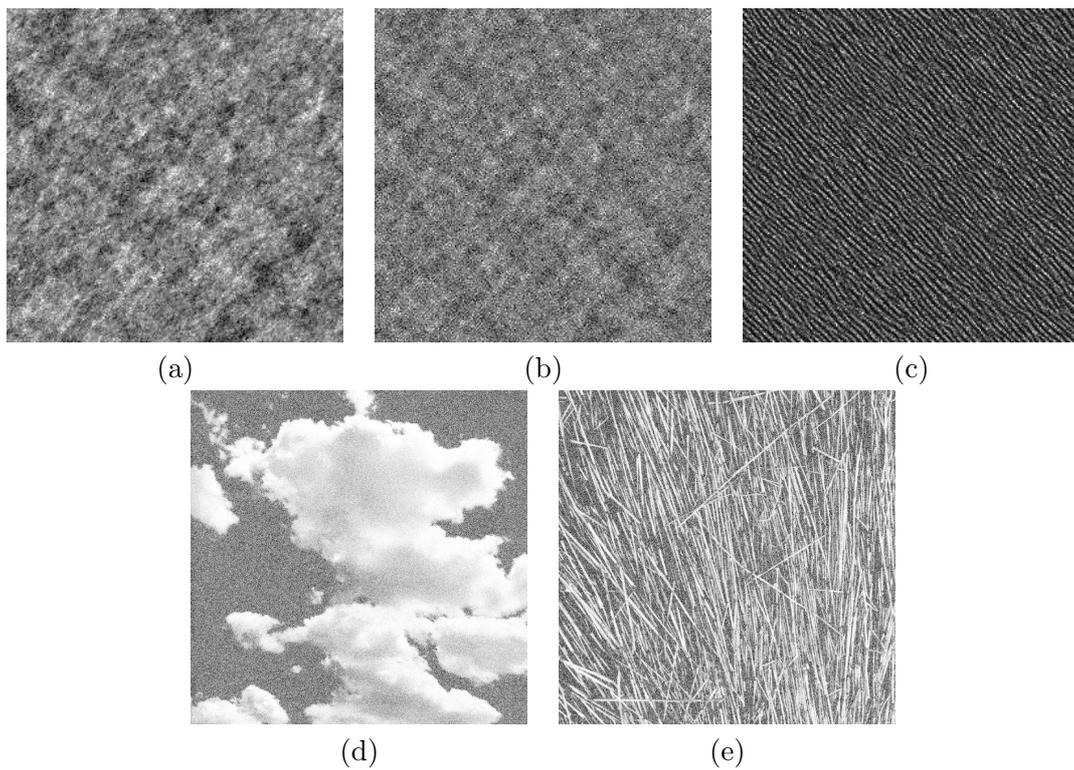


Figure 4.5: Speckled versions of the images in figure 4.4. The speckle noise has been generated with number of looks $L = 3$ according to (Rainey and Wessels, 1988): if L denotes the number of looks of the output image, L independent instances of Rayleigh distributed noise are generated, band-passed and averaged before being multiplied with the original noise-less image.

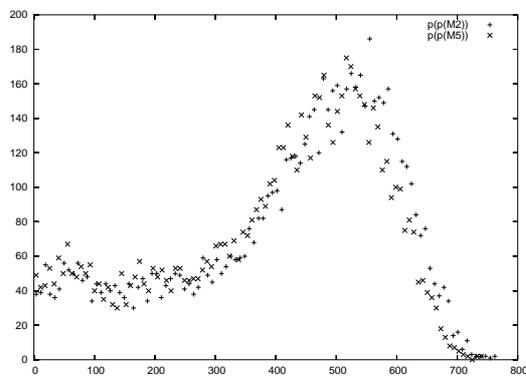
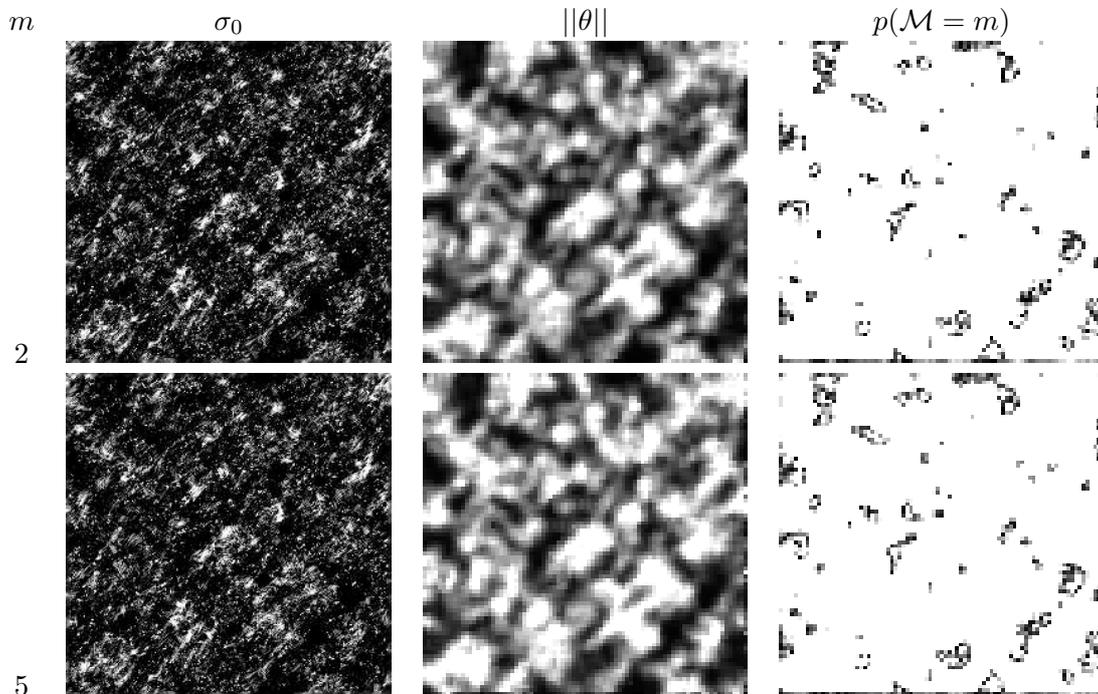


Figure 4.6: Denoising and information extraction with fixed model order — results on the synthetic, model order 2 noisy texture in 4.5a: for each model $\mathcal{M} = m$ used for processing, the model order, the obtained local backscatter σ_0 and texture $\|\theta\|$ intensities are reported together with the local evidence of the model $p(\mathcal{M} = m)$ and its histogram with “+” marking model order 2 and “×” marking model order 5. Although the two models considered are very similar to each other, model order 2 is most evident in the data, consistently with the simulation parameters.

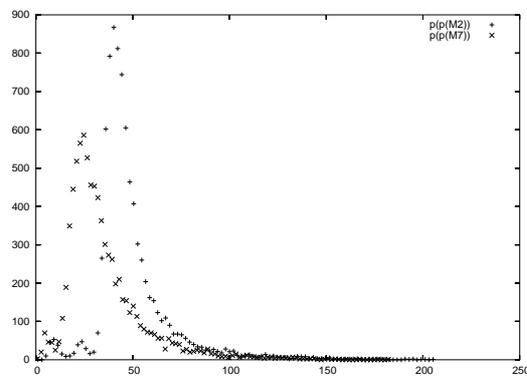
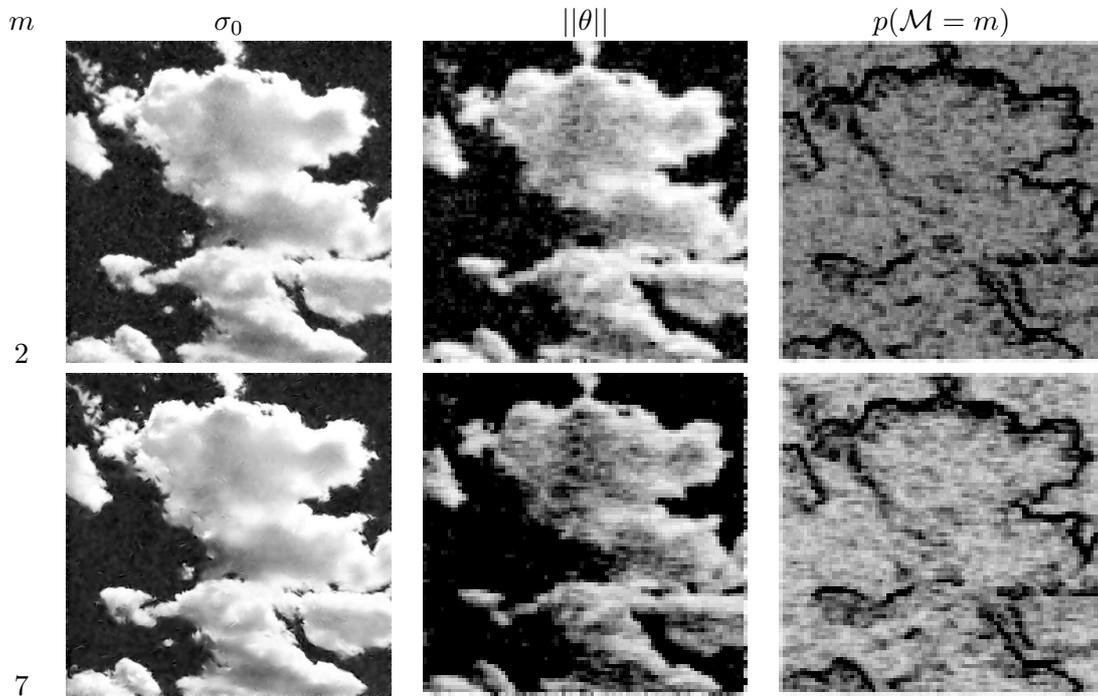


Figure 4.7: Denoising and information extraction with fixed model order — results on the natural Brodatz “cloud” image, with low geometrical complexity speckled texture in 4.5a: for each model m used for processing, the model order, the obtained local backscatter σ_0 and texture $\|\theta\|$ intensities are reported together with the local evidence of the model $p(\mathcal{M} = m)$ and its histogram with “+” marking model order 2 and “x” marking model order 7. Again, model order 2 is most evident in the data, consistently with the simplicity of the data.

noisy backscatter image $Y = y$

$$\Theta \rightarrow X \rightarrow Y$$

making explicit the model order m .

$$\mathcal{M} \rightarrow \Theta \rightarrow X \rightarrow Y.$$

The new hierarchical model is able to generate more diverse classes of realizations than the previous one: realizations with locally variable model complexity are now included in the model. This comes at the cost of a growing estimation complexity: the new causal relation must be inverted in addition to the ones related to the fixed order hierarchical model.

4.4.1 Variable order Gauss–Markov random field hierarchical modeling

As noted in section 3.2.6, Bayesian methods automatically and quantitatively include model order selection (Gull, 1988; Jeffreys, 1939). Complex models are automatically self-penalizing under Bayes’ rule.

The joint probability of a given model $\mathcal{M} = m$ and a given parameter $\Theta = \theta$ at a given location i in the image lattice with data pixel value $Y_i = y_i$ can be written explicitly as

$$\begin{aligned} p(\Theta_i = \theta_i, \mathcal{M}_i = m_i | Y_i = y_i) &= \frac{p(Y_i = y_i | \Theta_i = \theta_i, \mathcal{M}_i = m_i) p(\Theta_i = \theta_i, \mathcal{M}_i = m_i)}{p(Y_i = y_i)} \\ &\propto p(Y_i = y_i | \Theta_i = \theta_i, \mathcal{M}_i = m_i) \end{aligned}$$

by using the Bayes’ equation: the probability of the model is proportional to the evidence if a uniform non-informative prior for the location parameters $\mathcal{M} = m$ and $\Theta = \theta$ is assumed and if the data $Y = y$ is fixed. The evidence then naturally embodies Occam’s razor: in order to assign a preference to alternative models M_i , the first-level evidence $p(D = d | \mathcal{M} = m, \Theta = \theta)$ has to be evaluated.

If the despeckled backscattered pixel value $X_i = x_i$ is introduced as a nuisance parameter, the expression for the evidence of the model ($\Theta_i = \theta_i, \mathcal{M}_i = m_i$) in the data pixel becomes

$$p(\Theta_i = \theta_i, \mathcal{M}_i = m_i | Y_i = y_i) \propto \int_{X_i} p(Y_i = y_i | X_i = x_i) p(X_i = x_i | \Theta_i = \theta_i, \mathcal{M}_i = m_i) dx_i$$

where $X_i = [0, +\infty[$ is the space of the allowed values for the despeckled pixel values x_i .

If the integrand posterior distribution can be approximated by a Gaussian around its MAP maximum value — or, equivalently, if the monotonous logarithm of the posterior can be approximated by its second order Taylor expansion around the MAP (as shown for instance by Stan et al. (2001)), then

$$\begin{aligned} p(\Theta_i = \theta_i, \mathcal{M}_i = m_i | Y_i = y_i) &\simeq \int_X p(Y_i = y_i | X_i = \hat{x}_{i\text{MAP}}) p(X_i = \hat{x}_{i\text{MAP}} | \Theta_i = \theta_i, \mathcal{M}_i = m_i) \\ &\quad \cdot \exp\left(-\frac{1}{2} \Delta x_i^T H_i \Delta x_i\right) dx_i \\ &\simeq 2\pi p(Y_i = y_i | X_i = \hat{x}_{i\text{MAP}}) p(X_i = \hat{x}_{i\text{MAP}} | \Theta_i = \theta_i, \mathcal{M}_i = m_i) h_{ii}^{-1/2} \end{aligned}$$

with

$$h_{ii} \equiv -\frac{d^2}{dx_i^2} \left[p(Y_i = y_i | X_i = x_i) p(X_i = x_i | \Theta_i = \theta_i, \mathcal{M}_i = m_i) \right]_{|x_i = \hat{x}_{i\text{MAP}}}.$$

The log probability of the model given the data therefore becomes

$$\begin{aligned} \log p(\Theta_i = \theta_i, \mathcal{M}_i = m_i | Y_i = y_i) &\simeq \log p(Y_i = y_i | X_i = \hat{x}_{i\text{MAP}}) \\ &+ \log p(X_i = \hat{x}_{i\text{MAP}} | \Theta_i = \theta_i, \mathcal{M}_i = m_i) \\ &- \frac{1}{2} \log \left(-\frac{d^2}{dx_i^2} p(Y_i = y_i | X_i = x_i) p(X_i = x_i | \Theta_i = \theta_i, \mathcal{M}_i = m_i) \right) \end{aligned}$$

and by substituting Gauss–Markov prior (equation 4.1) and square–root Gamma likelihood (equation 4.2) in the expression:

$$\begin{aligned} &\simeq \left[(2L - 1)(\log y_i - \log \hat{x}_{i\text{MAP}}) - \log \hat{x}_{i\text{MAP}} - L \left(\frac{y_i}{\hat{x}_{i\text{MAP}}} \right)^2 \right] \\ &+ \frac{1}{2} \left[\log \sigma^2 - \frac{(\hat{x}_{i\text{MAP}} - \mu_i)^2}{2\sigma^2} \right] \\ &- \frac{1}{2} \left[-\frac{2L}{\hat{x}_{i\text{MAP}}^2} + \frac{6Ly_i^2}{\hat{x}_{i\text{MAP}}^4} + \frac{1}{\sigma^2} + \frac{1}{\sigma^2} \sum_{k,l}^{N_{m_i}} \theta_{k,l}^2 \right] \end{aligned}$$

where according to Walessa and Datcu (2000):

$$\begin{aligned} \hat{x}_{i\text{MAP}}^3 - \mu_i \hat{x}_{i\text{MAP}}^2 + L\sigma^2 \hat{x}_{i\text{MAP}} &= L\sigma^2 y_i \\ \mu_i &\equiv \sum_{k,l}^{N_{m_i}} \theta_{k,l} x_k. \end{aligned}$$

A contextual local model parameter and local model order estimation can be carried out by maximizing this obtained approximated expression for the probability $p(\Theta = \theta, \mathcal{M} = m | Y = y)$.

4.4.2 Local evidence maximization in Gauss–Markov random field image models

The selection of the most probable model order is performed by extending the Iterative Evidence Maximization loop to consider a spatially variant model order. The probability $p(Y = y | \Theta = \theta, \mathcal{M} = m)$ is maximized as a function of Θ and of \mathcal{M} .

The computational complexity grows with respect to the fixed model order case, since a new parameter has to be inverted by probability maximization.

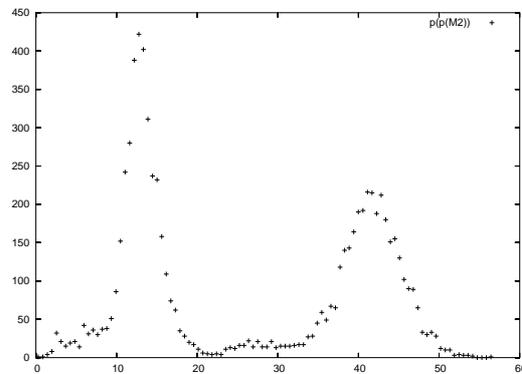
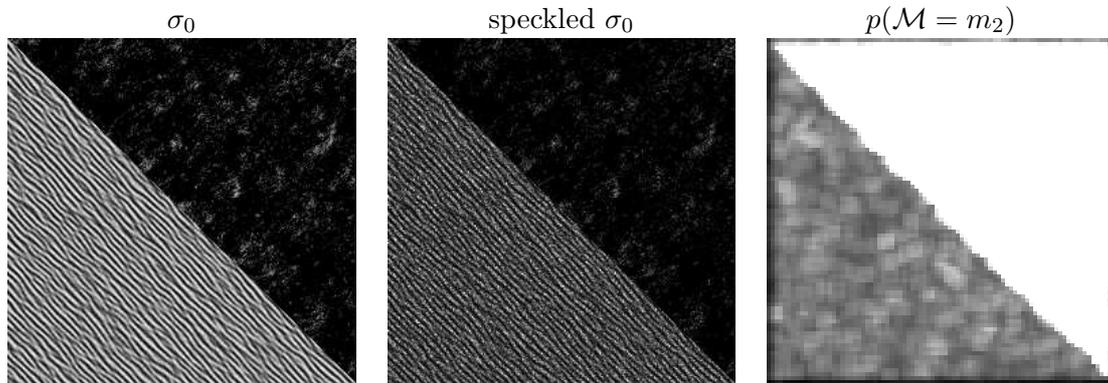


Figure 4.8: Clean and speckled multiple model (autobinomial and GMRF with order 2) test image, evidence of GMRF model with order 2 and histogram of this evidence. The multi-modality of the histogram – in abscissa, binned values of the model $M = m_2$ evidence $p(M = m_2)$ – is an indication that the models can be clearly separated by the evidence.

4.4.3 Model order selection examples and comments

Further “mixed” test images are produced by combining different samples of the considered models and processed, as in figures 4.8, 4.9 and 4.10, the multi-modality of the evidence histogram is an indication of the contextual presence of different models in the data. This indication is of course more pronounced in the cases in which the difference between models is more significant, but it is observable even in the case of GMRF models with different model orders.

The results in the estimated evidence and model order maps (in figure 4.11) tend to discriminate the different phenomena.

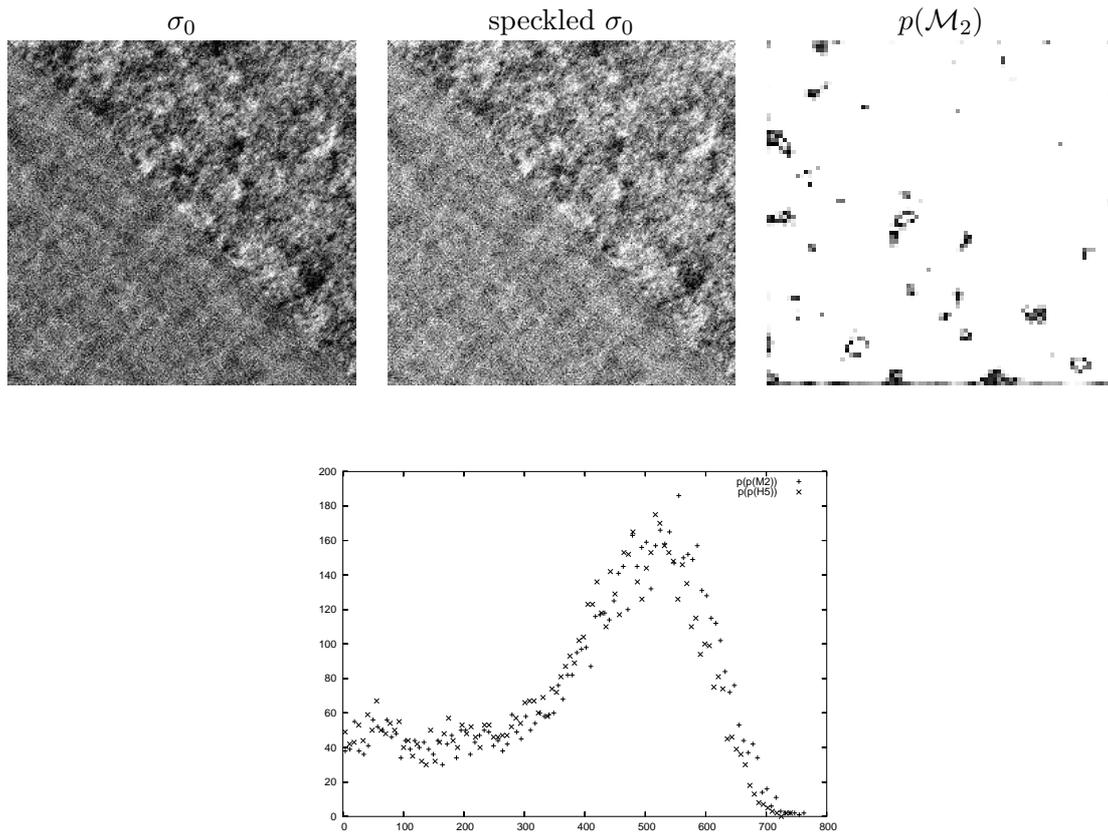


Figure 4.9: Clean and speckled multiple model (GMRF with orders 2 and 5) test image, evidence of GMRF model with order 2 and histogram of this evidence. In abscissa, again binned values for the model $M = m_2$ evidence $p(M = m_2)$. Two curves are clearly visible, but the discrimination between the two models in the evidence histogram is more problematic than in the case with different models.

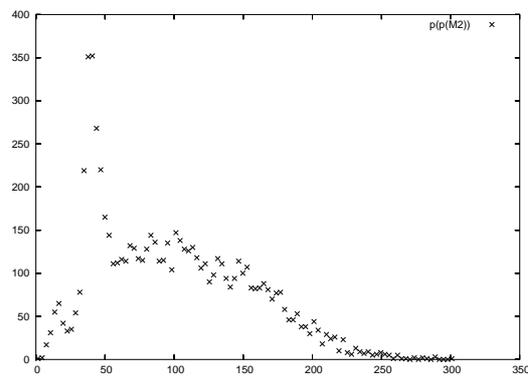
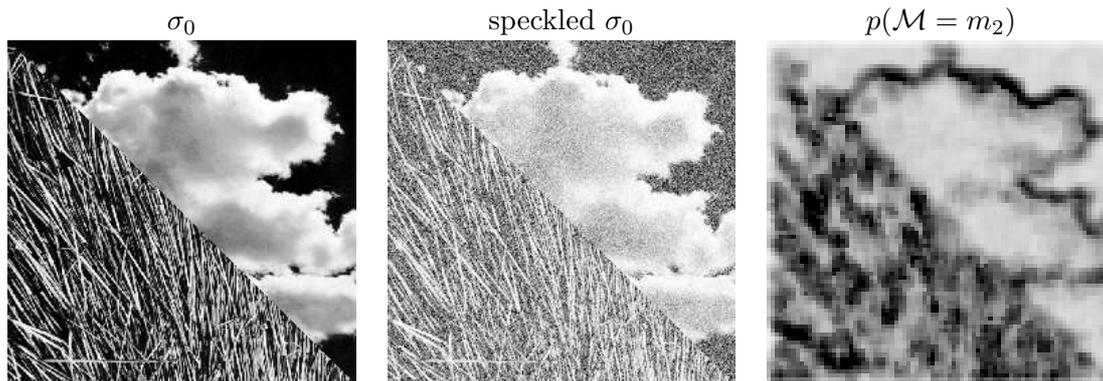


Figure 4.10: Clean and speckled multiple model (Brodatz “cloud” and “straw”) test image, evidence of GMRF model with order 2 and histogram of this evidence. In abscissa, binned values for the model $M = m_2$ evidence $p(M = m_2)$. Again, the clear bimodality suggests the contextual presence of multiple models in the data.

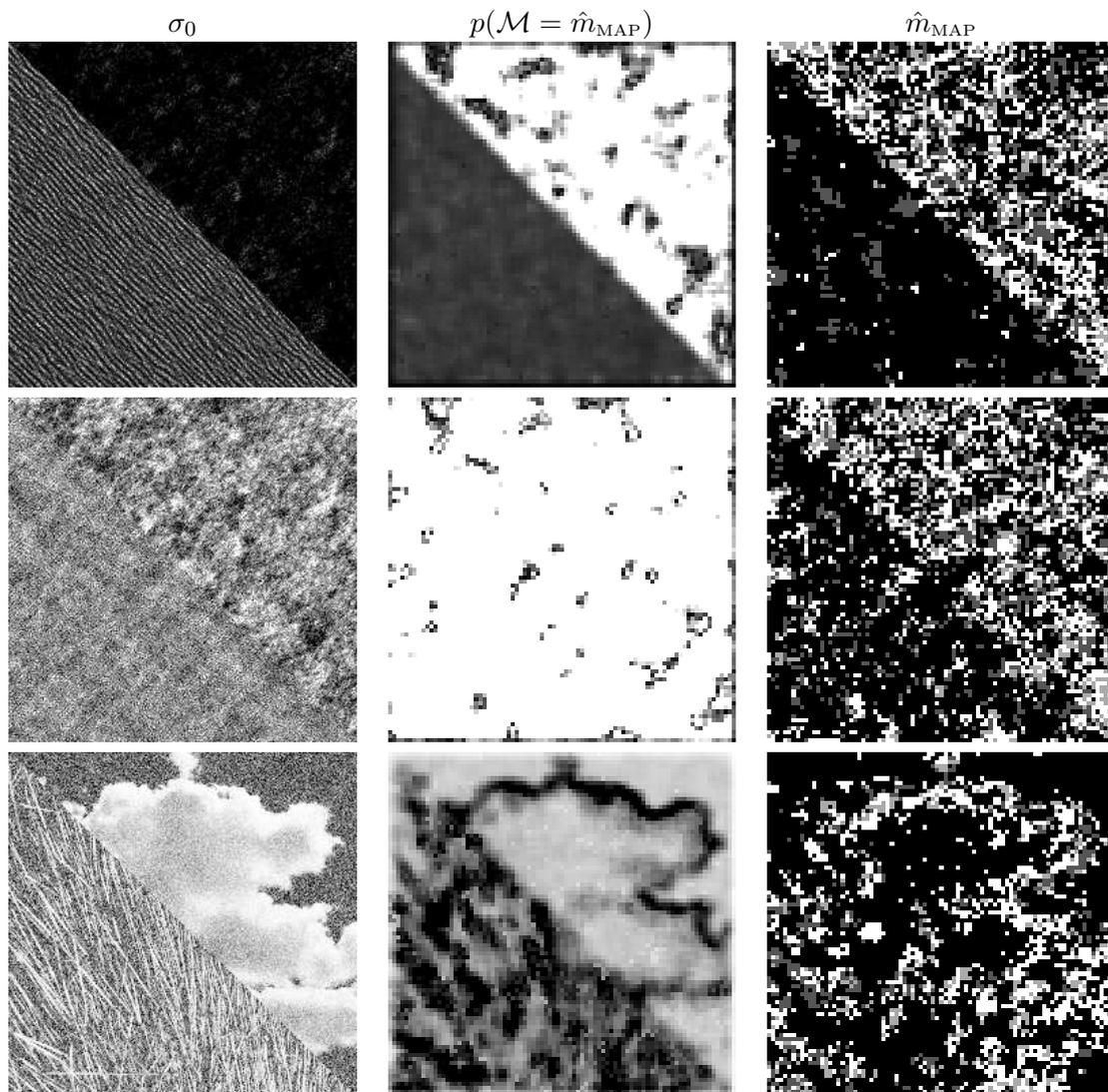


Figure 4.11: By using the second level of Bayesian inference based on the maximization of the local evidence across models, a local MAP estimate of the local model order can be performed from noisy data. The noisy intensity data, the evidence of the MAP model and the order of the MAP model itself are reported. In the right hand column, the discrimination of the order of the model from noisy data is performed even though the estimate itself is very noisy.

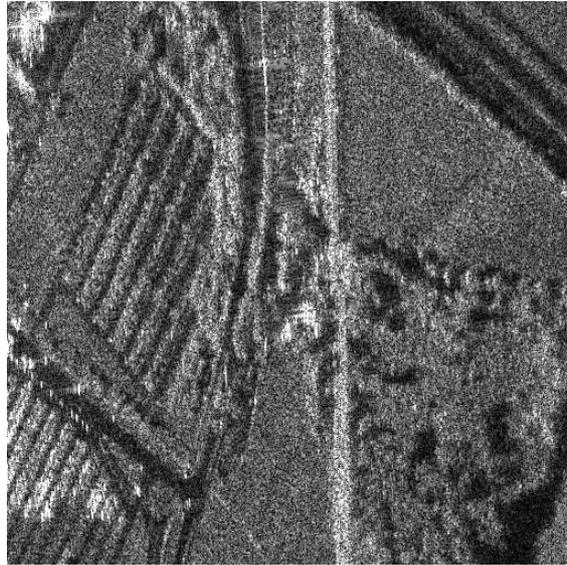


Figure 4.12: E-SAR X band SAR backscatter amplitude image over Oberpfaffenhofen, Germany.

4.4.4 Spatially-variant model order selection for SAR data analysis

We apply the extended image denoising and information extraction system to the analysis of an E-SAR X band SAR backscatter intensity image (in figure 4.12) acquired over the Oberpfaffenhofen site in Germany. The image contains elements of varying degrees of complexity including car parkings on the upper left, a forest on the lower right and a railway line cutting the scene vertically.

The processed data (in figure 4.13) show the quality of the despeckled dataset together with the locally estimated model order: a number of different structures in the data are recognized as being generated from processes of different complexity. The discrimination is able to distinguish between areas characterized by different geometrical complexity without taking into account the parameters of the local texture: smooth grassy areas are recognized as similar to shadow ones with respect to model complexity, just as forests and car parkings both share higher complexity.

4.5 Summary

In this chapter, an extension of the image denoising and information extraction method of Walessa and Datcu (2000) was introduced. The following points have been discussed:

- while the problematic of model selection in the Bayesian framework is well known, most applications in the image processing literature deal with clean data that are assumed completely known. Here instead the estimation is carried out locally starting from noisy data: the extended system incorporates a further estimation level that

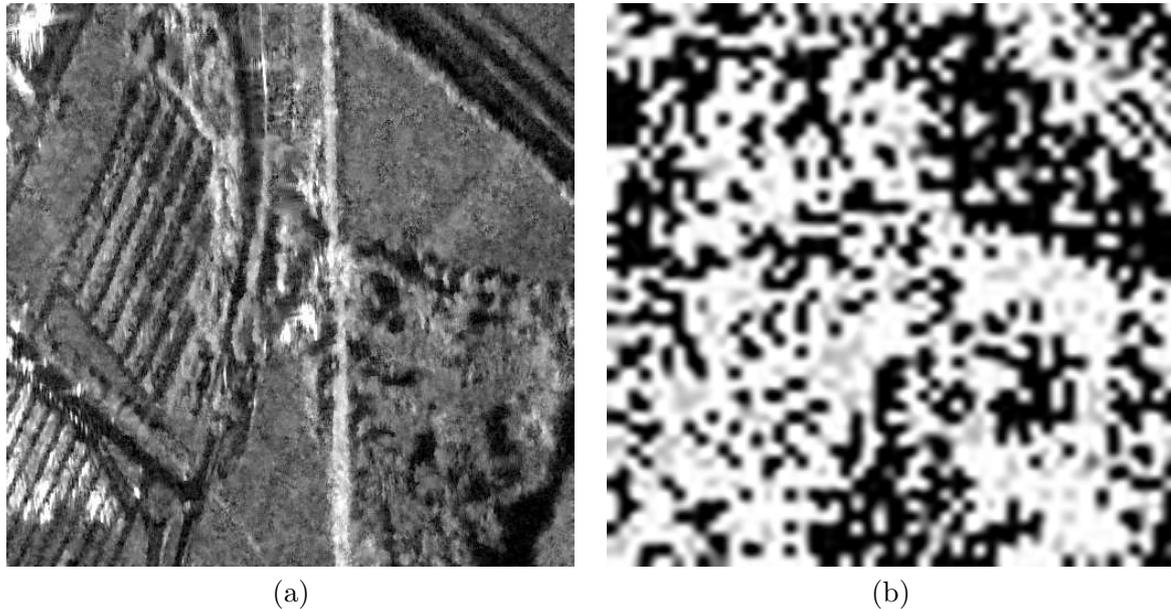


Figure 4.13: Despeckled image (a) and estimated model order (b) for the SAR intensity image in 4.12

allows the space-variant selection of the model-order by means of the maximization of the Bayesian model evidence;

- this extension allows the system to properly model and reconstruct images that are composed of instances of different stochastic image models with varying degrees of complexity;
- the extended algorithm was demonstrated on simulated data generated with well-known model orders as well as on Brodatz textures corrupted by multiplicative speckle noise;
- we gave examples of classification by model order in the context of Gauss-Markov random field model based image denoising and information extraction starting from speckled Synthetic Aperture Radar data. Model complexity maps were generated that demonstrate the discrimination potential of the Bayesian evidence maximization with respect to model choice.

The extended system is able to extract descriptors from the data in terms of a model-based vocabulary, to choose models for the data in a space-variant way in a given reference class and to discriminate different image elements by their complexity, again in terms of the model.

An application of the derived algorithm to the analysis of SAR data is presented in chapter III.

Chapter 5

Information Fusion for scene reconstruction from Interferometric SAR Data in Urban Environments

Abstract

A novel framework for scene understanding from InSAR data is presented that is based on Bayesian machine learning and information extraction and fusion.

The system is based on a hierarchical model of the acquisition process that leads from the scene to the data: this model is introduced and its different levels and components are detailed.

A generic description of the data in terms of multiple models is automatically generated from the original signals based on the denoising and information extraction algorithms presented in chapter 4.

The obtained feature space is then mapped by the application of simple Bayesian networks to user semantics representing urban scene elements in a supervised step.

The developed system is demonstrated by generating high-resolution land-use maps on urban environments from real InSAR data acquired from spaceborne sensors.

Interferometric SAR systems are able to contextually acquire multidimensional datasets whose channels, in particular the SAR backscatter intensity, the interferometric coherence and the interferometric phase, carry different kinds of information about the acquired scene:

- the phase is a very direct descriptor of the geometry of the scene in terms of the distances between the sensor and the scene elements (Massonnet, 1993);

- the interferometric coherence contains a description of the geometric and radiometric stability of the scene elements, and therefore allows a discrimination between different kinds of scene features;
- in the backscatter intensity, radiometric and geometric factors are again intermixed: many classes of scene elements are directly recognizable from their backscatter signature as well as from their spatial aspect.

To exploit the complementarity of the features derived from the different data sources, a Bayesian classification and information fusion framework can be employed that composes the different semantic-less descriptions of the data originated by the denoising and information extraction systems described in chapter 4 to associate them with user semantics mapping the acquisition into a well defined set of scene elements and features.

5.1 The need for probabilistic modeling in scene reconstruction

Scene reconstruction can be thought of as the procedure of finding the scene that best explains a given dataset:

$$\begin{array}{ll} \text{SCENE} \rightarrow \text{DATA} & \text{direct model of acquisition} \\ \text{SCENE} \leftarrow \text{DATA} & \text{scene reconstruction} \end{array} .$$

Examples of scene reconstruction include building recognition and the generation of land use maps from remote sensing data.

When the reconstruction is carried out from measurements derived from multiple sensors of dissimilar nature, or from sources that are fundamentally different — such as radar images and records of knowledge gathered from humans —, the techniques of information fusion are used to derive conclusions and decisions based on a synergy of all the data available.

A limitation of deterministic scene reconstruction algorithms is that the presence of uncertainties and noise in the available measurements — as well as the imperfect knowledge of the parameters of the acquisition — generates an inherent incompleteness of the data.

We therefore approach the problem of scene understanding and information fusion in a probabilistic framework.

In particular, we consider the nature of the scene reconstruction problem by adopting regularization strategies that incorporate external sources of information into the available data set.

Posterior probability as a measure of reconstruction quality In this context, the data d and the scene s are considered to be samples drawn from the multidimensional random processes D and S (Geman and Geman, 1984; Cressie, 1991).

A treatment of the scene reconstruction inversion problem that incorporates and composes models for these processes is obtained by rewriting the Bayes' equation as

$$p(S = s|D = d) p(D = d) = p(D = d|S = s) p(S = s) . \quad (5.1)$$

The main characteristic of the Bayesian approach is the explicitation of the modeling assumptions made for both the data formation mechanism $S \rightarrow D$ in the likelihood term $p(D = d|S = s)$ and for the phenomenon under analysis S in the prior $p(S = s)$. Further important features in this context include the capacity to deal with incomplete data, to learn causal relationships — thus learning the problem domain, and the power to combine knowledge and signal information.

The inversion can again be carried out by maximization of the posterior probability, resulting in the so called Maximum A Posteriori (MAP) estimate of the phenomenon S

$$\hat{s}_{MAP} = \arg \max_s \{p(D = d|S = s) p(S = s)\} . \quad (5.2)$$

Bayesian modeling and inversion as machine learning The MAP equation in (5.2) can be considered as a model for a learning device (Hackerman et al., 1995): a description for s is obtained from D by learning about their relation with the model S . Probability is used as a measure for the uncertainty in the knowledge of this relation.

5.2 Scene-to-data hierarchical models

When models for the scene S and the data D become so complicated that the inversion in terms of (5.2) becomes intractable, a common solution is a divide-and-conquer approach in which simpler intermediate description levels are introduced in the modeling.

Curse of dimensionality in stochastic data inversion The large amount of data required to perform a full inference while keeping high resolution detail observable would prohibit the treatment of high-order probability distribution functions. This explosion of computational complexity, known as the “curse of dimensionality”, has to be treated by recurring to models described by only local statistics, as is the case in Markov models and of hierarchical Bayesian networks.

5.2.1 Scene reconstruction Bayesian belief networks

A Bayesian network is set up for the task of scene reconstruction as in figure 5.1a. A tree edge structure encodes statistical independence between the random variables in the nodes. The last level of inference is performed by learning the problem semantics through user interaction.

Since the problem of inferring a general network structure from the data is in provably NP-hard, simplifying assumptions are often made about the nature of the existing dependencies between the variables.

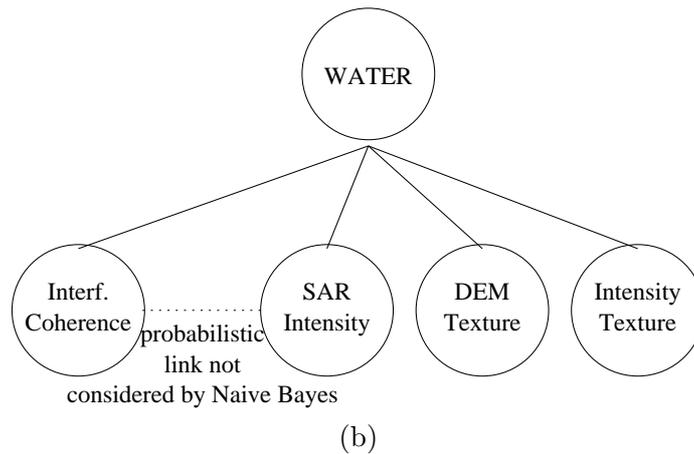
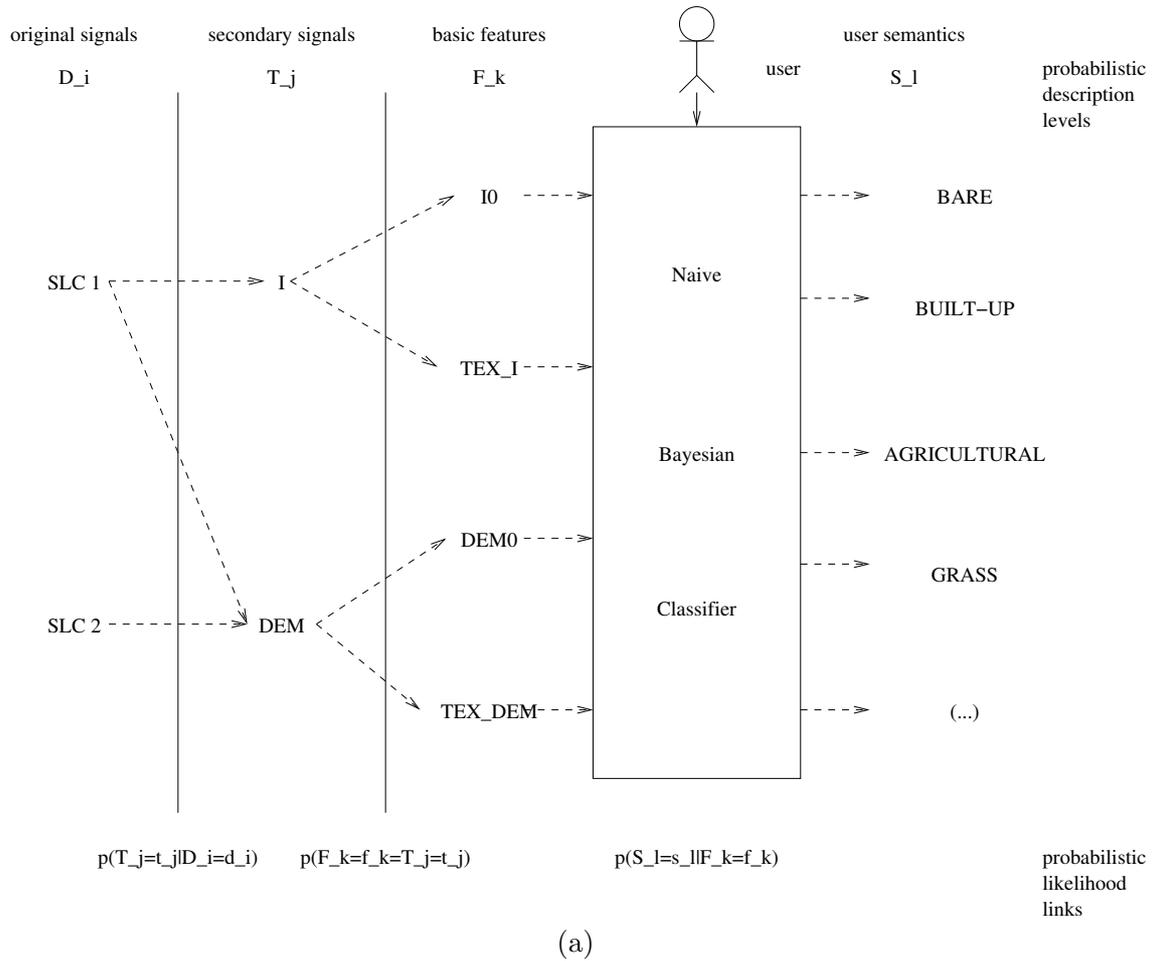


Figure 5.1: Learning scene semantics through hierarchical Bayesian Networks: (a) the tree edge structure encodes statistical independence between the random variables in the nodes. The last level of inference is performed by learning the problem semantics through user interaction. As in naive Bayesian classification variables on the same level of the tree structure are assumed mutually independent. This might not be strictly the case, as in the example in (b): the label WATER might be characterized by very low backscatter intensity and interferometric coherence. Although in general very low backscatter areas also show low coherence values, the relation between the two variables is not taken into account by the classifier.

As in naive Bayesian classifiers, the considered simplified network consist only of a series of levels of parents and of several mutually independent children nodes. Classification is obtained by considering the parent node to be a hidden variable stating which class each data element should belong to. The classification is termed “naive” since the hypothesis of statistical independence between variables belonging to the same level is not necessarily justified (see for instance figure 5.1b).

5.2.2 User feedback and machine learning

When the naive Bayesian approximation is considered valid, the scene S can be related to the data D through more levels of intermediate simpler models F_i which are causally linked

$$\begin{aligned} S \rightarrow F_1 \rightarrow \dots \rightarrow F_n \rightarrow D & \quad \text{direct modeling} \\ S \leftarrow F_1 \leftarrow \dots \leftarrow F_n \leftarrow D & \quad \text{scene reconstruction} \quad . \end{aligned}$$

The learning is in this case performed across levels, again by making use of the Bayes’ equation in a version that considers the various levels of inference involved

$$p(S = s|F_1 = f_1) \dots p(F_n = f_n|D = d) p(D = d) = p(D = d|F_n = f_n) \dots p(F_1 = f_1|S = s) p(S = s) \quad (5.3)$$

to derive a conclusion about the underlying scene S .

In this framework, a model for the image formation mechanism can be described in this way:

$$\begin{aligned} S \rightarrow \Omega \rightarrow R \rightarrow D & \quad \text{direct model} \\ S \leftarrow \Omega \leftarrow R \leftarrow D & \quad \text{scene reconstruction} \quad (5.4) \end{aligned}$$

a scene S is illuminated and generates a reflected field carrying information in the parameter ω , a realization of the scene feature process Ω . The scene features are translated to signal features R by the instrument and finally to data D by the image formation systems.

The scene reconstruction procedure must invert this causal chain, recovering the original scene features from the data. The layers of inference needed are implemented in the stages of a processing chain that is described in figure 5.2: the original signals acquired are used to generate image data products (backscatter intensity, coherence and height map) by using standard operational processors. These products are analyzed by a feature extraction system (described in chapter 4 and section 5.2.3) and translated into a set of image features. The features are grouped by similarity using a simple unsupervised classification system, and the generic multidimensional description obtained is fed into a Bayesian classification and information fusion system (section 5.3). This new level of description in the data-to-scene chain can incorporate user-provided semantic information into the available representation, finally connecting the generic, multiple-model based obtained description to the domain of the unknowns in the particular scene reconstruction problem under investigation.

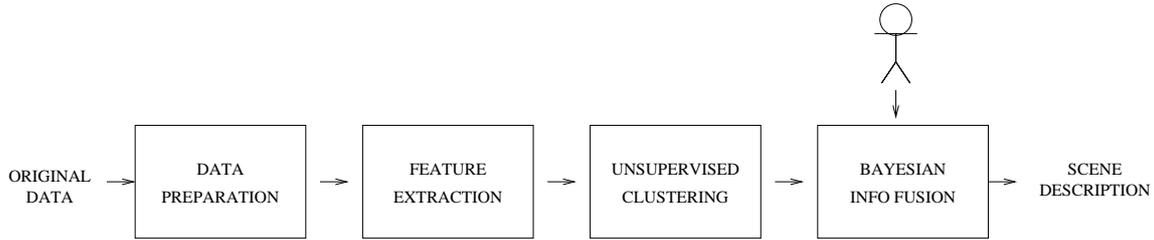


Figure 5.2: The scene reconstruction processing chain: data preparation is followed by information extraction. An unsupervised clustering module performs a grouping of the image features by similarity. Its output is subject to information fusion incorporating semantics by user feedback, thereby producing the final scene description.

5.2.3 Information extraction from backscatter and height image data

Since the height noise in interferometric SAR DEMs is additive Gaussian while the noise affecting SAR backscatter intensity is a form of strong multiplicative speckle noise, the feature extraction from SAR intensity and topographic height data can be performed respectively by the backscatter despeckling/Gaussian noise filtering and feature extraction systems introduced in chapter 4.

We postulate in this context that a quasi-complete description of a typical remote sensing image requires a separation of the sources of interesting information and of hindrances in the data — SAR backscatter despeckling, DEM noise filtering — and the estimation of the structural texture properties of the images thereby obtained.

The combined feature extraction system models and reconstructs a dataset that is free of noise, while still preserving its most important attributes: it estimates the parameters that describe the data in terms of the employed models. In this sense, it is not only a set of filters, but also a model-based feature extraction system that generates a representation of its input. The model fitting is performed in a space-variant manner that accommodates for local nonstationarities in the data.

5.3 Learning procedure

After the basic features have been extracted by the combined image information extraction system, they are grouped by similarity, by using a simple unsupervised K-Means classification system (appendix A). The multiple layers of information thereby obtained have then to be fused with each other: the clusters —here indicated as ω_i — play the role of an abstract image vocabulary that is able to explain, by different combinations, the image semantics S .

Instead of imposing an explicit definition of the phenomenon under study as in a rule-based expert system, descriptions are learnt from the human user by example. This implies a much more direct and powerful way of providing information to the system and enabling it to consider the interpreter conjectures.

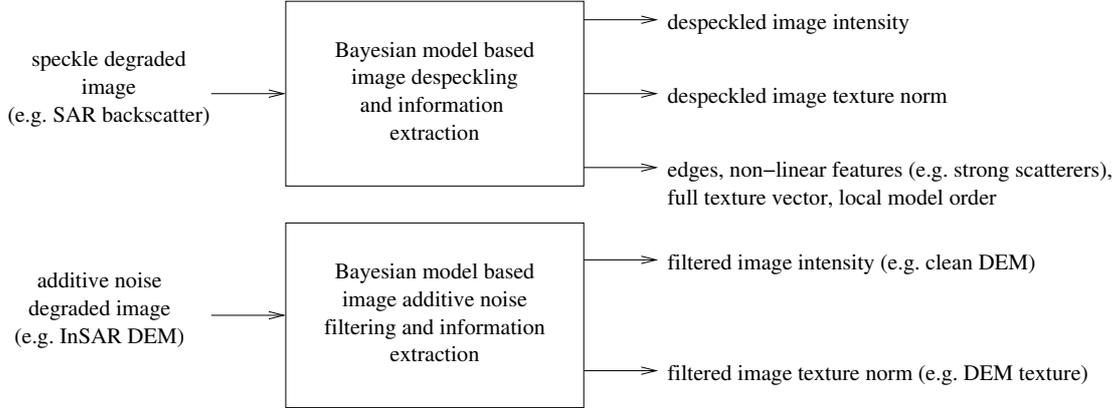


Figure 5.3: The model based SAR backscatter–InSAR derived height information extraction system: input square root intensity and interferometric height data are used to generate noise-clean estimates for the backscatter and for the height as well as local texture estimates.

Again, to be able to capture the subjective, interpreter-dependent aspects of information, a Bayesian formalization is needed: probability is interpreted as a degree of belief rather than as a frequency of realization. This contextual fusion and interactive classification is performed via a Bayesian classification and fusion system.

The process of interactive learning consists in progressively computing the probability of a specific cover-type in the image, on the basis of positive and negative pixel-level examples provided by the user. The inference process for the label $S = s_\nu$ given the image data d and the features in the clusters ω_i is realized through the probability

$$p(S_\nu = s_\nu | D = d) = p(S_\nu = s_\nu) \sum_i \frac{p(\Omega_i = \omega_i | S_\nu = s_\nu) p(\Omega_i = \omega_i | D = d)}{p(\Omega_i = \omega_i)}$$

where $p(S_\nu = s_\nu)$ and $p(\Omega_i = \omega_i)$ are (usually non-informative) prior probabilities, while $p(\Omega_i = \omega_i | S_\nu = s_\nu)$ has to be learnt from user examples. An independence condition is assumed, i.e., the probability is assumed equal to the product of the separate likelihoods for the cover type d_ν given each single model

$$\begin{aligned} p(\Omega_i = \omega_i | S_\nu = s_\nu) &= p(\Omega_{i,M_1} = \omega_{i,M_1}, \Omega_{i,M_2} = \omega_{i,M_2}, \dots, \Omega_{i,M_m} = \omega_{i,M_m} | S_\nu) \\ &= p(\Omega_{i,M_1} = \omega_{i,M_1} | S_\nu = s_\nu) p(\Omega_{i,M_2} = \omega_{i,M_2} | S_\nu = s_\nu) \dots \\ &\quad p(\Omega_{i,M_m} = \omega_{i,M_m} | S_\nu = s_\nu) \end{aligned}$$

since each cover type is considered a combination of different models

$$\omega_i = \omega_{M_1} \otimes \omega_{M_2} \otimes \dots \otimes \omega_{M_m}$$

where \otimes denotes scalar product between sets.

If we denote with ϕ the vector parameter which satisfies the identities

$$p(\Omega_i = \omega_i | S_\nu = s_\nu, \Phi = \phi) = \phi_i, \quad p(\Phi = \phi) = \Gamma(r)$$

where $\Gamma(\cdot)$ is the Gamma distribution, then the probability of the training set is

$$p(\Phi = \phi|T) = \text{Dir}(\phi|1 + N_1, \dots, 1 + N_r)$$

where $\text{Dir}(\cdot)$ is the Dirichlet distribution and N_i is the number of occurrences of the signal type ω_i in T . If a new training set T' is provided, the probability is updated according to

$$p(\Phi = \phi|T, T') = \text{Dir}(\phi|1 + N_1 + N'_1, \dots, 1 + N_r + N'_r)$$

Denoting with a the parameter vector which satisfies the identity

$$a_i = 1 + N_i$$

the learning process is modeled by updating the vector a after observation of each training set (Schröder et al., 2000a).

5.4 Urban scene reconstruction from InSAR data

We apply the described framework to the analysis of urban scene data acquired by airborne SAR sensors with metric resolution.

5.4.1 E-SAR Dresden data analysis for settlement understanding

A single polarization high resolution interferometric E-SAR dataset on the German city of Dresden is analyzed. The city center lies in the area on the lower left of the image. The Semper-Oper building belongs to the area.

False color interferometric land use color composites Laur et al. (1998) as the ones in figures 5.4b and 5.6b are useful to allow an easier discrimination between different land use types: mostly green areas correspond to heavily vegetated (forests) or layover areas, blue areas to water surfaces (sea and inland water), red areas mostly to bare rock and stable agricultural fields and mostly yellow ones to urban centers.

Starting from the interferometric dataset, the SAR backscatter denoising and information extraction system extracts despeckled backscatter and full texture vector from the $|I_1|$ master amplitude. We also extract the interferometric phase. After K-means unsupervised classification, we obtain vectorized data for these three quantities.

The intensity clustering produces quite clear results, while the phase information might be complicated by unwrapping errors.

Texture strength and variance are extracted from the texture vector. The obtained channels are then separately subject to K-means clustering to produce a classified feature space representation.

A supervised classification and fusion is carried out starting from these preclassified datasets. The supervised MAP classification/fusion obtained tends to discriminate between the different kinds of ground cover and land use present in the imaged scene: natural

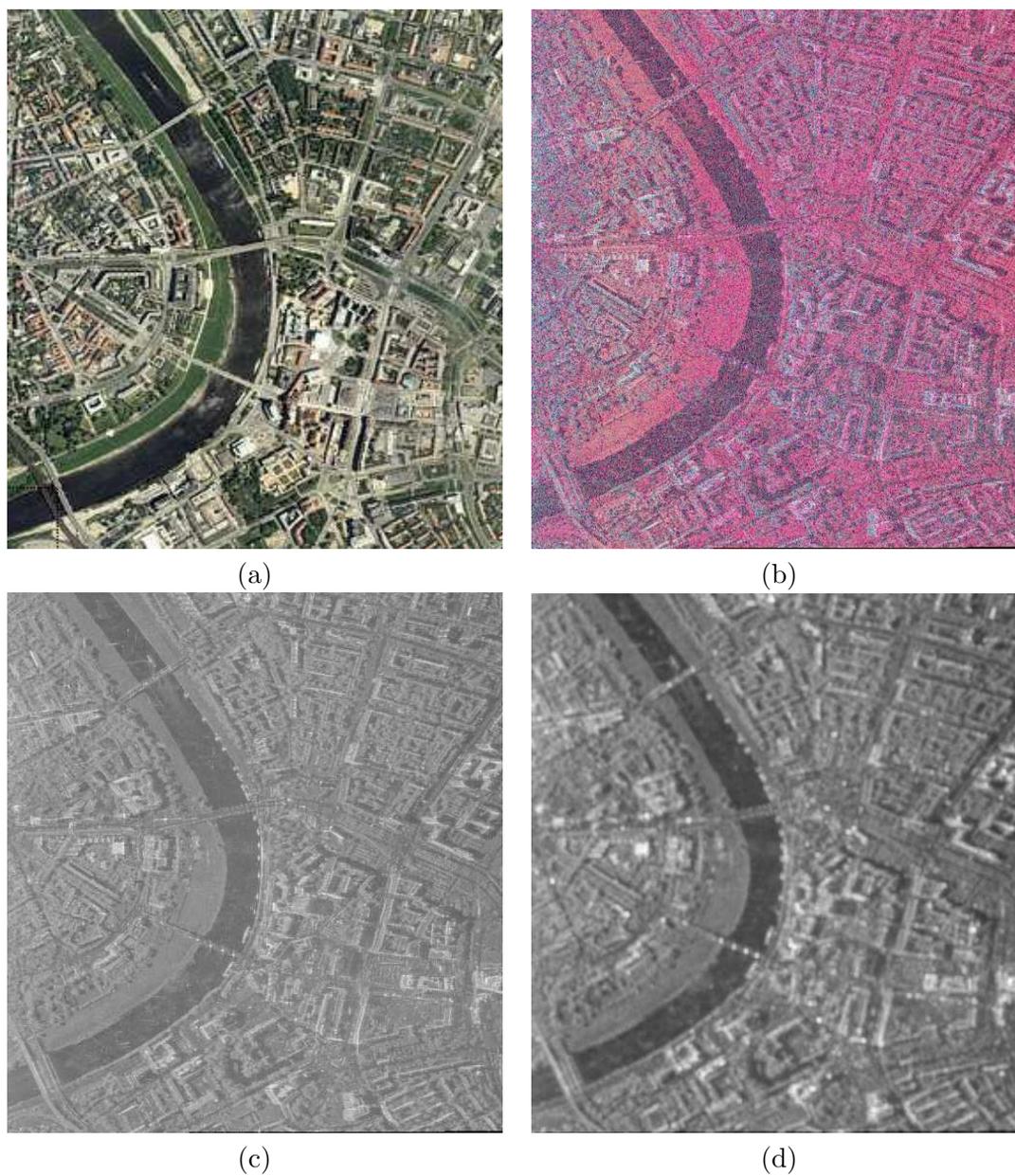


Figure 5.4: E-SAR Dresden - Ground truth optical image (a), interferometric land use (Laur et al., 1998) false color image (R=coherence, G=average intensity, B=intensity difference) (b), SAR backscatter intensity (c), SAR texture norm (d).

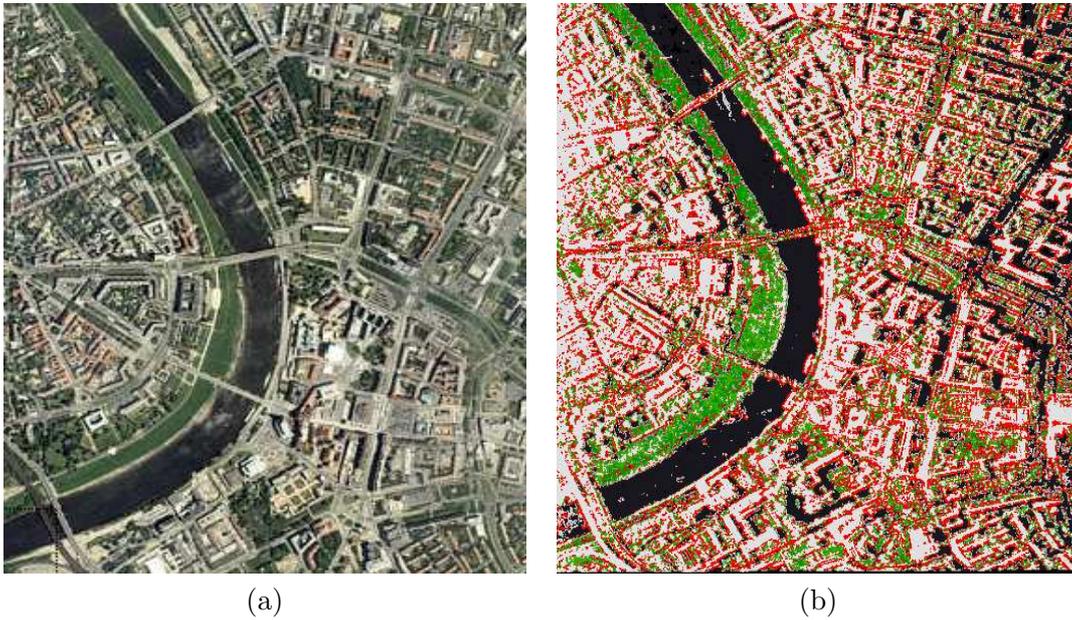


Figure 5.5: E-SAR Dresden - Ground truth optical image (a) and generated land use map (b) with classes vegetation (GREEN), vertical flat surface (WHITE), strong scatterer (RED), water/shadow (BLACK).

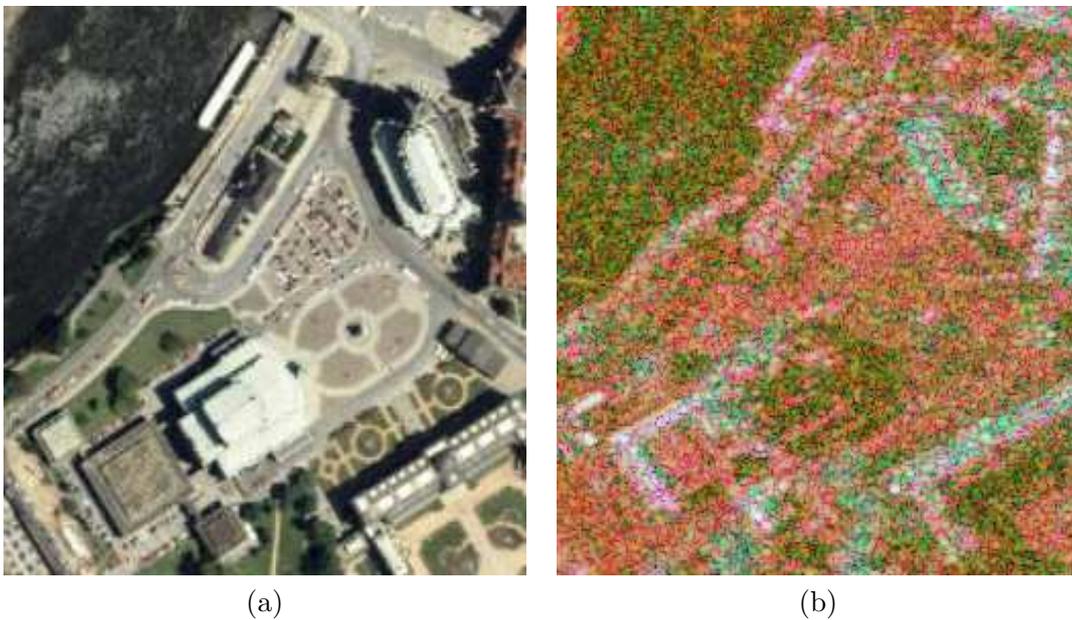


Figure 5.6: E-SAR Dresden Semper Oper- Ground truth optical image (a), interferometric land use (Laur et al., 1998) false color image (R=coherence, G=average intensity, B=intensity difference) (b).

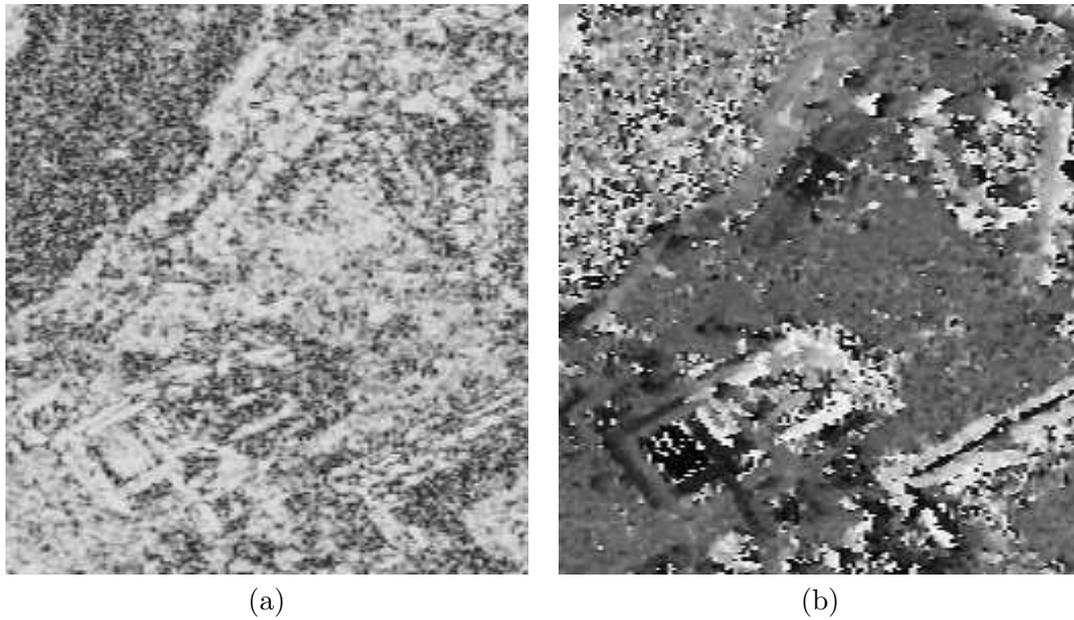


Figure 5.7: E-SAR Dresden Semper Oper- interferometric SAR backscatter coherence (a), interferometric SAR phase (b).

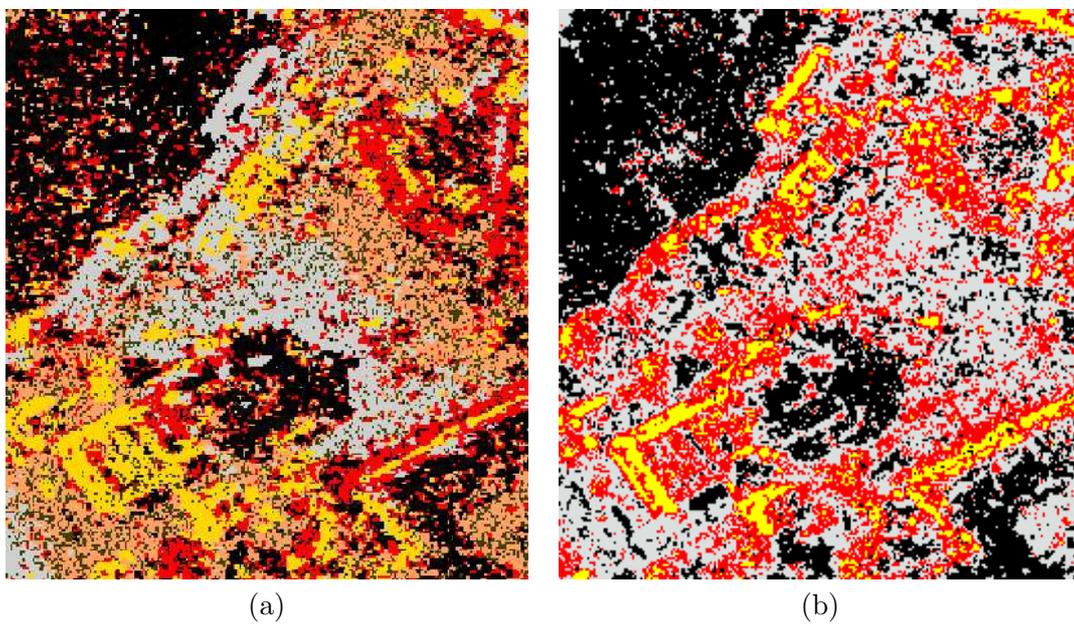


Figure 5.8: E-SAR Dresden Semper Oper- land use map obtained by unsupervised classification (a) and Bayesian classification/fusion (b) by considering intensity, coherence, absolute phase: vertical surfaces in layover (YELLOW), strong scattering areas (RED), flat terrain (LIGHT GRAY), layover areas (ORANGE), water/shadow (BLACK).

(grass, water, trees) as well as man-made scene elements (walls, roofs) are extracted and separated from each other yielding results that outperform those obtained by simple K-means clustering conducted across all available channels.

5.5 Summary

In this chapter, a novel framework for scene understanding from InSAR data was presented that is based on Bayesian machine learning and information extraction and fusion. The following points were detailed:

- the system is based on a hierarchical model of the acquisition process that leads from the scene to the data: a generic description of the data in terms of multiple models is automatically generated from the original signals, based on modifications of the denoising and information extraction algorithms presented in chapter 4;
- unsupervised K-means clustering is used to group by similarity the elements populating the feature space;
- the obtained feature space is then mapped by the application of simple Bayesian networks to user semantics representing urban scene elements in a supervised step.

The developed system was demonstrated by generating high-resolution land-use maps on urban environments from real InSAR data acquired from spaceborne sensors. Further applications of the method to the reconstruction of urban scenes from high-resolution InSAR data are reported in chapter III.

Chapter 6

Stochastic geometrical modeling for urban scene understanding from a single SAR intensity image with meter resolution

Abstract

To overcome the limitations of pixel-based analysis applied to sub-metric resolution data, a novel model based algorithm for the automatic reconstruction of building areas from single observation meter-resolution Synthetic Aperture Radar intensity data is introduced.

The reconstruction is based on the Maximum A Posteriori estimation by Monte Carlo methods of an optimal scene that is modeled as a set of pairwise-interacting Poisson-distributed marked points describing parametric buildings.

Each of the objects can be hierarchically decomposed into a collection of radiometrically and geometrically specified object facets that in turn get mapped into data features.

The geometric and radiometric phenomenologies of sub-metric SAR in urban environments are taken into account in the likelihood term.

The detection of the buildings is based on a decision likelihood ratio.

The optimization of the obtained posterior is based on Gibbs sampling and Monte Carlo Markov chains.

The developed algorithm is demonstrated on real data acquired by airborne X band sensors on built-up areas in high complexity settings characterized by tall vegetation and strong metallic scatterers.

The Bayesian classification and information fusion approach to scene reconstruction presented in the last chapter is applicable in the range of resolutions from tens of meters down

to about one meter. At lower scales, two phenomena take place in urban environments:

- information in SAR *data* becomes inherently sparse, in the sense that the areas of the image containing some energy (mostly corresponding to cavities and sharp edges) are intermixed by much larger areas of very low SNR caused by the specular reflection of the incident signal on smooth surfaces: a traditional lattice-based model should consider an inordinately large amount of (mostly totally noisy) pixels in order to gain some discrimination and modeling capability;
- a pixel-based description of the *scene* would anyway be of limited interest, since at sub-metric resolutions single objects are recognizable as separate entities associated with specific semantics and generating peculiar effects in the data.

The usual interferometric approach based on unwrapping tends to be difficult in high resolution SAR data, since very steep surfaces and areas of very low coherence are the norm. Unwrapping needs to be regularized by means of external information. If, for instance, the tower in Figure 6.1 is considered, it is evident that extreme layover, multiple reflections and the presence of very strong, high coherence echoes in combination with very low energy, low coherence returns from non-dihedral scattering will render the usual lattice-based interferometric reconstruction approach quite limited.

The artifacts complicating the analysis are actually due to the geometry of the scene: this is why they carry information about it. We therefore concentrate our analysis on the case of single pass, single model high resolution SAR data on urban environments. Phenomena usually considered as limitations of the dataset and excluded from further processing are instead recognized and exploited as information sources to provide added robustness to the estimation algorithm.

Normal lattice-based image data models and inversion algorithms can then be effectively complemented by descriptions in terms of scene objects. Evident features in the data have to be exploited in the scene reconstruction problem. Hindrances (e.g. a dihedral scattering from a building base partially hidden by the surrounding vegetation) can be overcome by using a prior description for the scene. A hierarchical modeling of remote sensing data can be constructed that considers the whole structure from data to reconstructed scene objects.

6.1 Scene understanding from SAR: objects and their relations

Scene understanding (SU) is in Computer Vision the recognition and reconstruction of the objects in the scene and of the spatial relations between them (Edelman and Weinshall, 1989). The problem is related to that of understanding visual perception in information processing terms: assuming that the purpose of vision is to reconstruct the spatial layout of the outside world, one is confronted with the problem of ambiguity in inferring the third dimension from retinal projection. This ambiguity is a necessary consequence of the imaging process, during which the depth information is lost. Marr (1982) proposed to compensate

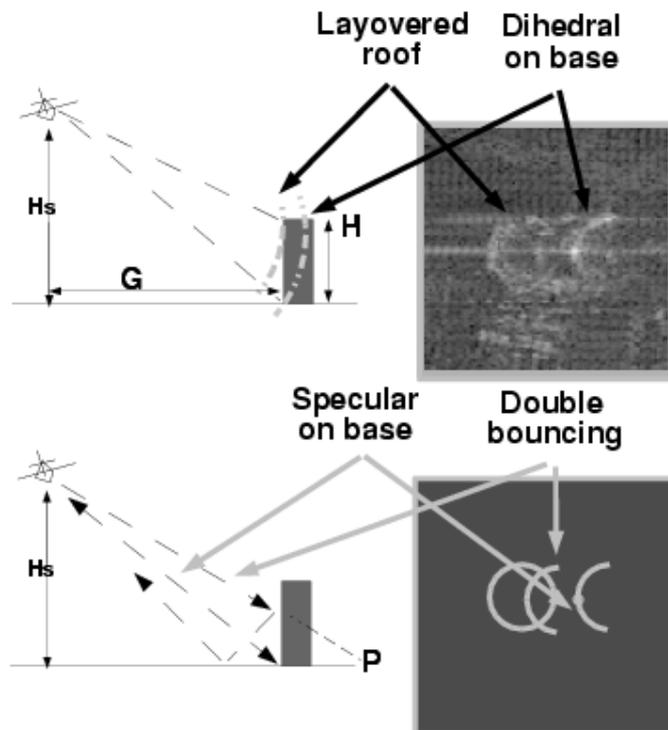
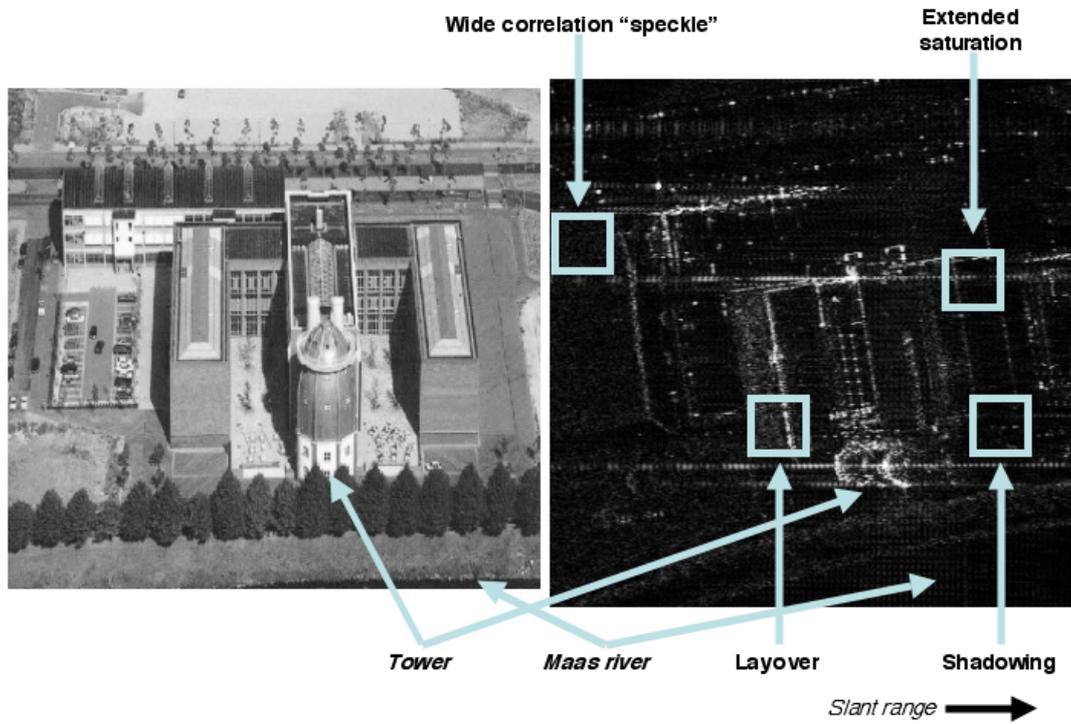


Figure 6.1: Meter resolution interferometric SAR phenomenology ground-to-slant mapping: basic signal features. Two distinct high energy, high interferometric coherence returns from roof (layovered) and base characterize buildings in meter resolution SAR imagery.

for this loss by constraining the solution to the reconstruction problem to conform to a priori assumptions, dictated by our knowledge of the physical world.

As the resolution of SAR systems approaches the level where the scene can be represented as a set of separated interacting objects, the methods and techniques of computer vision, traditionally centered on optical data taken on man-made settings, can be employed in radar remote sensing inversion problems, provided proper modeling of the specificities of the sensor and of the scene is assured.

6.2 Marked spatial point processes as prior scene models

A marked spatial point process (Upton and Fingleton, 1985; Diggle, 1983; Cressie, 1991) on the space X with marks in \mathcal{F} can be represented mathematically as a point process on the product space $X \times \mathcal{F}$. It can be defined in terms of random locations of events in a set X and corresponding random marks $Z(\cdot)$ in a set \mathcal{F} and used to model random events occurring at sparse, random locations according to its marginal spatial bidimensional point pattern. For example, events might represent the unknown positions of buildings in a scene and the corresponding marks may be describing them parametrically.

Gibbs point processes In particular, Gibbs marked point processes of order m are a class of inhomogeneous point processes for which the intensity function is written in terms of exponentials of functions of up to m distinct point process events. If $m = 2$, the probability distribution can be written as

$$p(S = (s_1, \dots, s_n), N = n) = \frac{\exp(-\nu(A))}{\alpha n!} \exp\left[-\sum_{i=1}^n g_1(s_i) - \sum_{i=1, j=1, j \neq i}^n g_{1,2}(s_i, s_j)\right]$$

with $g_1(\cdot)$ and $g_{1,2}(\cdot, \cdot)$ two positive first- and second-order potential terms, ν and α two free parameters. We approach the scene understanding problem by considering a marked point process to model built-up areas: a random parametric description $O(\cdot)$ of the geometry of a building object marks a random site C_i representing its geometric position (Ortner et al., 2002) in a study area D imaged in the data:

$$S \equiv \{(C_i, O(C_i)), C_i \in D\}.$$

The expression for the posterior corresponds to that for a generic Markov point process, in which the joint density function $p(S)$ can be factored uniquely into a series of terms if only the first two are considered.

$$p(S, n) = \frac{1}{Z(n)} \exp\{-U_n(s_1, \dots, s_n)\}$$

$$U_n(s_1, \dots, s_n) = \sum_{i=1}^n g_1(s_i) + \sum_{1 \leq i < j \leq n} g_{1,2}(s_i, s_j) + \dots$$

with U_n the total potential energy of the Markov point process.

A Gibbs marked point process can be demonstrated to be equivalent to a Markov point process. The spatial point process that we consider is Markov with range ρ ,

$$p(O(C_0)|S \setminus O(C_0)) \equiv p(O(C_0)|\{O(C)\}, C \in b(C_0, \rho) \setminus \{C_0\}) \quad (6.1)$$

with \setminus indicating the set subtraction operation, where $b(C_0, \rho)$ is a ball of center C_0 and ray ρ in the domain D : scene objects tend to interact with each other to produce more plausible ordered configurations. The properties of the marginal spatial point process (and in particular its density) are conserved in the marked point process, except for a finite number of sites in some vicinity of each of the given points in which local (repulsive) interactions are modeled that tend to decrease the density of the original marginal point process.

6.2.1 Point and mark distributions

The building centers are Poisson-distributed in the domain D . The marks describing the building geometry are instead defined as in

$$O(C) \equiv (W, \alpha, Q)$$

with uniformly distributed random parameters

$$Q \text{ one of } \begin{cases} \text{cylinder} \\ \text{block} \\ \text{v-shaped roof} \end{cases} \quad \begin{array}{l} \text{building shape element} \\ \\ \end{array}$$

$$W \in \mathbb{R}^3 \quad \text{width, length, height of building element}$$

$$\alpha \in [-\pi/2, +\pi/2] \quad \text{base angle to ref. direction.}$$

6.2.2 Modelling hierarchy

We further assume that each of the separate building objects O_i in which we can decompose the scene S in turn can be described as composed by different facets F_{ij} . Each of those can be mapped to a set of SAR amplitude pixels with values A_k in the SAR amplitude data image space (Figure 6.2): the parameters of each building define together with the acquisition geometry its observable facets and the pixels that belong to them

$$S \equiv \{O_i\}_i \equiv \{\{F_{ij}\}_j\}_i \rightarrow \{A_k\}_k \equiv D .$$

The hierarchical structure models a container/contained relation whose top level represents the full scene while the bottom one is a set of data pixels.

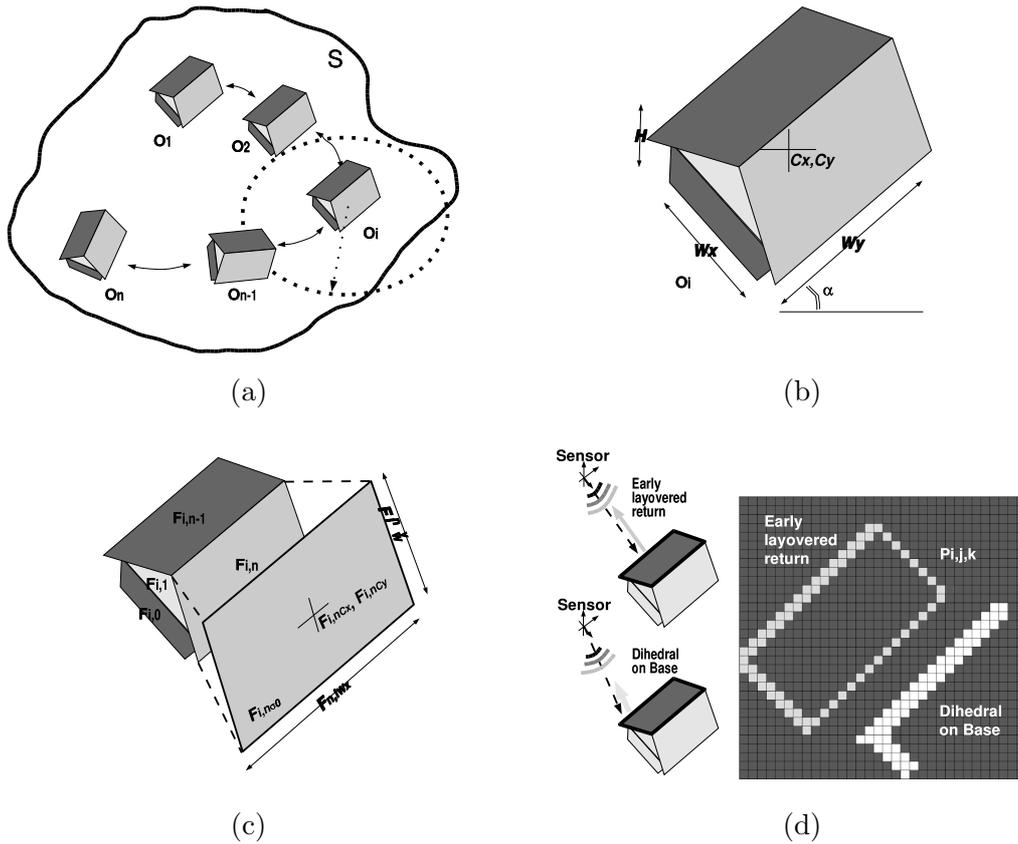


Figure 6.2: Levels in the hierarchical scene model: (a) the scene S is a collection of buildings O_i that are pairwise interacting within a given range; (b) objects are described parametrically as marked Poisson-distributed points; (c) objects can be decomposed into a collection of facets F_{ij} that have both a geometric and a radiometric parametric probabilistic description; (d) the facet F_{ij} is probabilistically mapped into a set of pixels $\{D_k\}_k$ in the amplitude image data space.

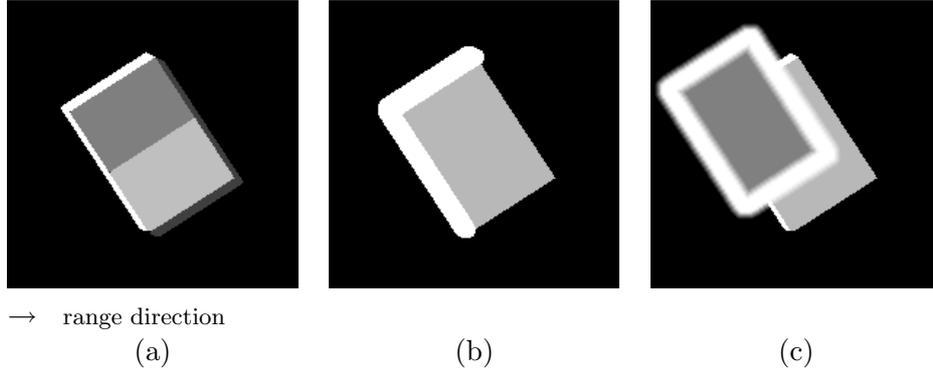


Figure 6.3: SAR phenomenology considered in the scene understanding process: (a) for every scene object (in this case a v-shaped roof house) (b) dihedral scattering region and (c) layovered roof are considered in the computation of the likelihood potential as clues for the reconstruction.

6.3 Scene posterior structure decomposition in Gibbs potential terms

Given a single meter resolution SAR intensity image, we want to infer by Maximum A Posteriori (MAP) estimation a description of the scene that is most likely to have generated it: if S and D denote respectively the scene and the data stochastic processes, then $p(S = s|D = d)$, the posterior probability we maximize, can be understood as a measure of quality for the reconstructed scene given the data.

Scene posterior description in terms of Gibbs formalism The posterior probability can be written under a Gibbs field formalism (Winkler, 1995) as in

$$\hat{s}_{\text{MAP}} = \arg \max_S p(S = s|D = d) = \arg \max_S \frac{1}{Z} \cdot \exp[-U(s|d)]$$

where Z is a normalizing constant (the partition function) and $U(s|d)$ is a local potential function. $U(S = s|D = d)$ can be decomposed into a regularization prior potential $U_P(S = s)$ while a likelihood potential $U_L(D = d|S = s)$ describes a distance between scene and data that takes into account a model of the acquisition system:

$$U(S = s|D = d) \sim U_P(S = s) + U_L(D = d|S = s).$$

Scene posterior decomposition The potential terms can therefore be decomposed as in

$$U_P(S = s) = \sum_{o_i, o_j \in s, j \neq i} U_p(O_i = o_i|O_j = o_j)$$

$$U_L(D = d|S = s) = \sum_{o_i \in s} \sum_{F_{ij} = f_{ij} \in o_i} \sum_{a_k \in d} U_l(A_k = a_k|F_{ij} = f_{ij})$$

where $U_p(\cdot)$ and $U_l(\cdot)$ are the object-level prior and the pixel-level likelihood potentials and where some of the possible dependencies have been neglected to simplify the modeling in the spirit of naive Bayesian classification: when the descriptions for both the scene and the data formation mechanism become too complex, both theoretical considerations connected to the “curse of dimensionality” (Jaynes, 1987) and practical concerns about the tractability of the model suggest the adoption of a divide-and-conquer approach in which simpler intermediate representation levels are introduced in the modeling. In our case, for simplicity, elements at the same level of the description are considered independent of each other, as in the case of Naive Bayesian Classification. This assumption allows a much simpler handling of the estimation process.

6.4 The scene prior potential term

To keep the modeling as simple and general as possible, the potential term

$$U_P(S = s) = \sum_{i,j} U_p(O_i = o_i, O_j = o_j)$$

is kept totally independent of the data and the sensor characteristics, and therefore describes an a priori plausibility of the scene only.

Keeping the prior term as general as possible, and therefore avoiding introducing local maxima in the posterior distribution through it, is desirable both to keep the reconstruction algorithm general and to ease the subsequent search for the global MAP. The prior interaction potential $U_p(\cdot)$ is used to avoid overlapping conditions between scene objects and corresponds to a geometric intersection measure between different scene objects:

$$U_p(O_i = o_i | O_j = o_j) \sim \text{area of intersection}(o_i, o_j) .$$

6.5 The data likelihood potential term

While the prior potential is independent of the data, the likelihood potential must instead express a definition of the relevant acquisition phenomenology. Both the geometric and radiometric causal relations that compose the direct model of the acquisition are expressed through the likelihood.

6.5.1 Extraction of strong targets from clutter in urban environments by inverse Gaussian statistics, decision theory-based classification

Strong signals generated by dihedral scattering on the building basement and by direct layovered returns from clean edges on the building roofs need to be modeled and extracted to be exploited in $U_l(\cdot)$.

Inverse Gaussian distributed backscatter values (Section 2.3.2) are assumed to model the amplitude return from the slowly varying small number of scatterers that compose these

elements: they have been shown to approximate well a K-distribution while keeping a better tractability, and have been used to model backscatter in urban environments (Frery et al., 1997; Mejail et al., 2001).

An expression for the discriminating likelihood ratio

$$\frac{p(A_k = a_k | \alpha_s, \gamma_s, L)}{p(A_k = a_k | \alpha_0, \gamma_0, L)} \quad (6.2)$$

with

$$p(A_k = a_k | \alpha, \gamma, L) = \exp \left[-\frac{L^L \Gamma(L - \alpha) a_k^{L-1}}{\gamma^\alpha \Gamma(-\alpha) (\gamma + L a_k)^{L-\alpha}} \right]$$

is obtained with $L > 0$ the number of looks of the data, $a_k > 0$ the amplitude pixel value, and $-\alpha_s, \gamma_s > 0$ and $-\alpha_0, \gamma_0 > 0$ respectively the parameters for the distribution that models the amplitude values of the strong scattering facet elements, and those for remaining clutter obtained by solving the system

$$\begin{aligned} \gamma &= L \left[\frac{\hat{m}_1 \Gamma(-\alpha) \Gamma(L)}{\Gamma(-\alpha - 1/2) \Gamma(L + 1/2)} \right]^2 \\ \frac{\hat{m}_{1/2}^2}{\hat{m}_1} &= \frac{\Gamma^2(-\alpha - 1/4) \Gamma^2(L + 1/4)}{\Gamma(-\alpha - 1/2) \Gamma(L + 1/2) \Gamma(-\alpha) \Gamma(L)} \end{aligned}$$

where \hat{m}_n is the n -th order moment estimated on a training area and $\Gamma(\cdot)$ is the Gamma function.

6.5.1.1 Radiometric likelihood potential term

Based on the discriminating likelihood in Equation 6.2, the likelihood potential can be written as

$$U_l(A_k = a_k | F_{ij} = f_{ij}) = \begin{cases} \frac{p(A_k = a_k | \alpha_s, \gamma_s, L)}{p(A_k = a_k | \alpha_0, \gamma_0, L)} & \text{if } a_k \text{ taken in } f_{ij} \\ 0 & \text{otherwise} \end{cases}$$

$$p(A_k = a_k | \alpha, \gamma, L) = \exp \left[-\frac{L^L \Gamma(L - \alpha) a_k^{L-1}}{\gamma^\alpha \Gamma(-\alpha) (\gamma + L a_k)^{L-\alpha}} \right].$$

Given a hypothesis on the 3D disposition of strong scatterers in the scene, the loci of the strong scattering phenomena can be defined in deterministic terms, as in Figure 6.1. Since the likelihood ratio potential $U_l(A_k = a_k | F_{ij} = f_{ij})$ is non-zero for the candidate facet pixel values only, its computation requires the determination of the geometry of the strong responses from the visible facets in the data: the candidate scene object edges are iterated upon, they are projected in slant range coordinates and the loci in the image of the relevant dihedral scattering and roof layover phenomena to be analyzed are therefore determined.

The expression for the decision ratio in the last paragraph provides by itself a discrimination between strong clutter and backscatter that are typical of very high resolution SAR data.

It can therefore be directly used to assess the probability that a candidate strong scattering area in the scene space (linked to some phenomenology in the image space by a ground-to-slant transformation) appears as such in the data.

6.5.2 Geometric likelihood potential terms

The analysis of the data must take into account the peculiar aspects of the phenomenology of high resolution radar. Strong responses often appear translated in range due to layover (and to multiple signal reflections), but no full geocoding is possible, since local heights at metric resolutions are unknown.

In particular, if H_s is the altitude of flight of the sensor with respect to the ground, $H = h$ is the building height and G its position in ground range, then the layovered signal generated from building roof edges is located at a slant-range distance Δ_{rg} of

$$\Delta_{rg} = (H_s^2 + G^2)^{1/2} + ((H_s - h)^2 + G^2)^{1/2}$$

from the first one produced by dihedral scattering at the border between the vertical structure and the terrain.

The loci of the hypothesized top and bottom of the scene building facets are therefore used as the loci on which the estimation of the likelihood potential is performed.

6.6 Inference in Bayesian scene understanding with non-analytic posteriors

MAP inference is performed globally on the whole scene parameter space. The practical computation of the posterior probability $p(S = s|D = d)$ of a given candidate scene S is performed by decomposing it in terms of the hierarchically dependent elements O_i and F_{ij} that constitute it. The α_0, γ_0 and α_s, γ_s factors are estimated starting from clutter and strong scattering training areas. A decision likelihood ratio image can then be generated from the data and used directly in the computation of the U_l likelihood potential.

The obtained posterior expression is not analytic: as a consequence, the MAP estimate of the optimal scene given the data

$$\hat{s}_{\text{MAP}} = \arg \max_s p(S = s|D = d)$$

cannot be computed directly. The optimization is instead carried out by Monte Carlo methods.

6.6.1 Monte Carlo Markov Chain Gibbs sampling of the scene posterior

We adopt a simulated annealing procedure (Metropolis et al., 1953) for the estimation of the MAP (see Appendix B) : the multivariate posterior is maximized by iteratively

evaluating many possible solutions in a multidimensional parameter space. In stages, the optimization algorithm restricts the search paths to only the most promising solutions. While at the beginning of the iterative optimization procedure the sampled posterior function is essentially flat, at the end of the annealing the algorithm is forced to follow a steepest ascent trajectory. The net effect is that of maximizing the $f(\cdot)^{1/T}$ for T that slowly changes from high values to zero to be able to find the maximum of the function.

Scene configuration acceptance ratio and cooling schedule An acceptance ratio that is dependent on the rate of improvement in the solution is used together with the decreasing “temperature” T acceptance factor to probabilistically accept or reject new solutions. The Metropolis algorithm can therefore be understood as a relaxation of the greedy maximal ascent algorithm.

This approach allows the system to explore solutions that a direct maximum-gradient maximization would not consider, permitting it to identify the global maximum of a multivariate distribution.

The acceptance “temperature” factor is decreased from starting temperature T_0 with iteration step $iter$ according to a logarithmic cooling schedule as in Stoica et al. (2000)

$$\tau(i) = T_0 / \log(1 + 2\sqrt{iter})$$

in order to allow for a sufficiently slow confinement of the optimization algorithm to steepest ascent paths.

6.6.2 Scene configuration initialization

For simplicity, the initialization starting solution that is input into the system at temperature T_0 consists of a set of identical objects, all located at the same location in the center of the image and in the minimum allowed number of objects. More sophisticated starting configuration schemes (such as the initialization of the system by some simpler sub-optimal scene understanding algorithm) might improve the convergence speed.

6.6.3 Scene configuration sampling and Gibbs dynamics

A Gibbs sampler is set up to sample the posterior distribution keeping all the dimensions of the scene process fixed but one: this modified distribution can be shown (Winkler, 1995; Stoica et al., 2000) to converge to the full posterior.

For generating the new candidate solutions to the optimization problem from the previous ones, update rules are defined as in (Ortner et al., 2002) with the addition of the mutation of the building type parameter Q .

The scene marked point process with varying number of objects is simulated under the spatial birth-and-death process framework, as shown by Imberty and Descombes (2002). Special care must be used in ensuring that the mutation dynamics of the Gibbs sampler

are completely reversible, to avoid getting stuck on some local maxima or on the borders of the domain. Each of the defined mutation step must have an equiprobable opposite move. The moves considered are object creation and destruction, and displacements and mutation moves for the building parameters.

All the transformations are symmetric: first a building object

$$U \equiv (C_i, O(C_i)) = u \equiv ((x, y), (\alpha, w, q))$$

centered at location $C_i \equiv (x, y)$ in a lattice and having parameters α (the orientation angle), w (a three-terms width, length and height size term) and q (the building element type) is uniformly chosen in the image spatial domain, and then the mutation is applied to it:

- Translations are generated by

- uniformly sampling $\delta = (\delta_x, \delta_y)$ from $[-d_x, +d_x] \times [-d_y, +d_y]$ and

- setting

$$u = ((x, y), (\alpha, w, q)) \rightarrow v = ((x + \delta_x, y + \delta_y), (\alpha, w, q))$$

considering for this transformation that the image space is a torus, in order to obtain the symmetry of the transformation.

- Rotations are generated by

- sampling δ_α uniformly in $[-d_\alpha, +d_\alpha]$, and

- setting

$$u = ((x, y), (\alpha, w, q)) \rightarrow v = ((x, y), (\alpha + \delta_\alpha, w, q)).$$

In particular, we consider $[-d_\alpha, +d_\alpha] = [-\pi/2, +\pi/2]$.

- Dilations are generated according to the same principle: the length, width or hight of the object are modified. This modification is applied contextually to the dimension of the object and to its center coordinate, in order to let one side of it constant under the transformation, as shown in Figure 6.4, by

- uniformly sampling $(\delta_{DX}, \delta_{DY}, \delta_{DZ})$ from $[-d_{DX}, +d_{DX}] \times [-d_{DY}, +d_{DY}] \times [-d_{DZ}, +d_{DZ}]$ and

- setting $u = ((x, y), \alpha, w, q)$ to

- * $v = ((x + \delta_{DX}/2 \cdot \cos \alpha, y + \delta_{DY}/2 \cdot \sin \alpha), (\alpha, w + (\delta_{DX}/2, \delta_{DY}/2, \delta_{DZ}), q))$
for rectangle-shaped objects and

- * $v = ((x + \delta_{DX}/2, y + \delta_{DY}/2), (\alpha, w + (\delta_{DX}/2, \delta_{DY}/2, \delta_{DZ}), q)$ for round shaped ones for which the components of $W = w$ are by definition coincident.

We notice that this kind of dilation transformation is the one that is naturally applied to height mutations of buildings on a plane: the roof of the building is moved, while the its base remains fixed.

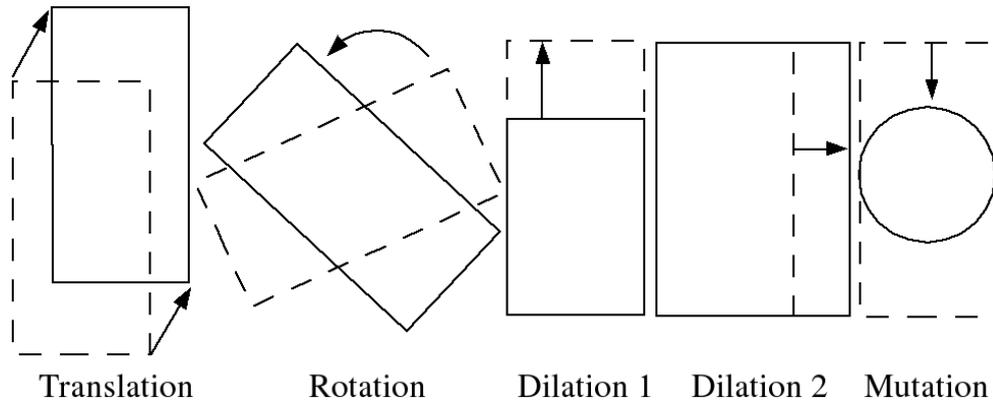


Figure 6.4: Gibbs sampler dynamics. Dilation moves are preferred to width and length modification with fixed center, since they leave one of the borders of the object totally untouched.

- Mutations are generated by sampling a q' uniformly in the space $\{\text{cylinder, box, v-shaped roof building}\} \setminus q$ (with \setminus again the set exclusion operator) and consequently setting

$$u = ((x, y), (\alpha, w, h, q)) \rightarrow v = ((x, y), (\alpha, w, h, q'))$$

with the convention that for transitions to a cylinder from square-based objects only the first element of w is kept as a radius of the new base, whereas for transitions from round objects to square ones the new horizontal components of w are again decoupled but keep at the next step after the mutation their common value equal to the cylinder radius.

For simplicity, the initialization starting solution that is input into the system at temperature T_0 consists of a set of identical objects, all located at the same location in the center of the image and in the minimum allowed number of objects. More sophisticated starting configuration schemes (such as the initialization of the system by some simpler sub-optimal scene understanding algorithm) might improve the convergence speed.

6.7 Overview of estimation method

The full estimation system is described in Figure 6.5.

Simulated annealing is often considered too computationally expensive to be used in image processing. Since the scene understanding system considers only a limited number of objects in the scene compared to the huge number of image lattice sites, the dimensionality of the problem is reduced, speeding up convergence. The only operations performed at every iteration are the mutation of one of the parameters of the model and then, given the updated scene model, the recomputation of the mutated potential terms in the posterior.

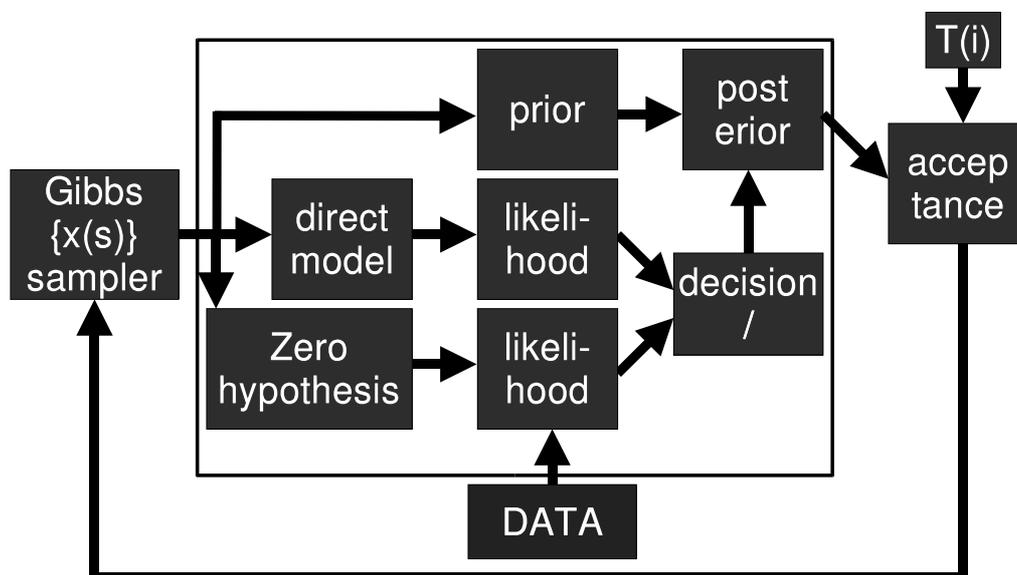


Figure 6.5: The complete MAP estimation system: simulated annealing is used for the maximization of the posterior via Gibbs sampling. The posterior depends on an a priori defined measure of plausibility of the hypothesis scene and on a likelihood ratio for the actual building detection.

In the implementation, all objects provide iterators to their component sub-objects. In particular, Facet objects provide iterator access to facet edges, corners and, by using a simple acquisition geometry model, to data pixels.

While the prior is computed at the object level by evaluating a geometrical intersection measure between different scene objects (O'Rourke, 1998), the computation of the likelihood potential needs to take into account the data pixel-level description of the buildings in the candidate scene (Simonetto, 2002; Haines, 1994) and the likelihood Ratio image.

When the MAP optimization is carried out iteratively, only the terms related to mutated objects need to be recalculated at each iteration.

The quantities related to the non-mutated scene elements (the intersection potential values and the data related ones that are generated by taking into account pixel-level descriptions of the buildings that are dynamically generated from the object parameters) are constant and are therefore not object of re-computation, yielding better performance.

The optimization procedure on images containing about a million of pixels takes about 30 minutes on a common PC workstation.

6.8 Marked Point Process based scene understanding example

In Figure 6.6, Intermap AeS-2 data on Trudering test area, east of Munich. Some houses are visible (a), together with the generated shadows and a part of the road network. In (b), the main identified buildings are shown with their location, size and orientation. The limitation of the obtain results is explained by the complexity of the scene: the structures to be reconstructed are small houses partially hidden at X band by surrounding taller vegetation. Typically for very high-resolution SAR data in built-up environments, multiple scattering and other disturbing effects tend to dominate the data, saturating the receiver and propagating to different areas in the image.

6.9 Summary

In this chapter, a novel algorithm based on stochastic geometry, decomposable object models and Gibbs sampling was introduced. The following points were detailed:

- an algorithm for the reconstruction of building areas from a single high resolution SAR intensity image was introduced, described and demonstrated on data with resolutions ranging from 0.5 to 2 meters;
- the algorithm is based on the Maximum A Posteriori estimation via Monte Carlo methods of an optimal scene described in a hierarchical Bayesian framework as a set of pairwise interacting Poisson distributed marked points that represent parametrically described scene buildings;

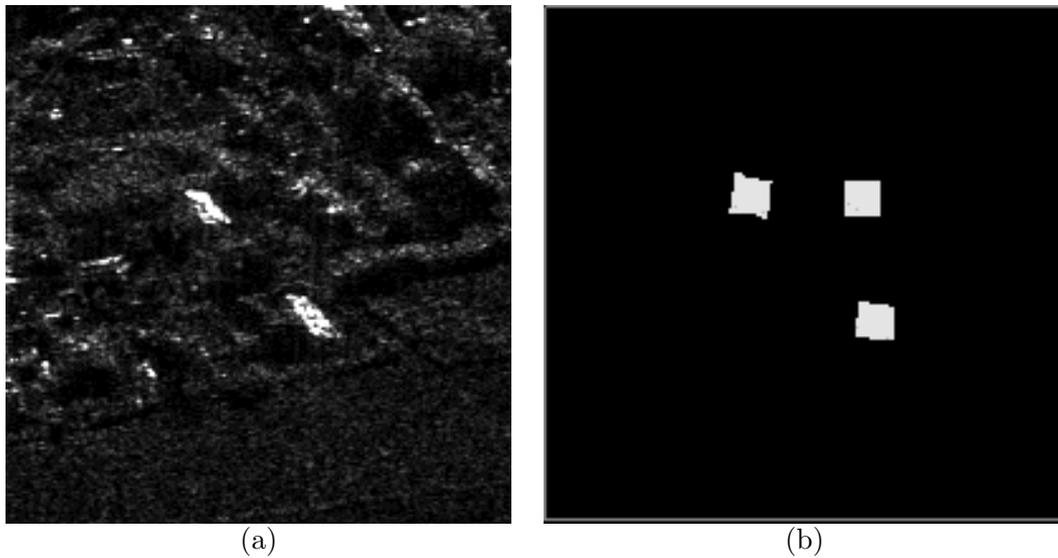


Figure 6.6: Intermap AeS-2 data on Trudering test area, east of Munich. Some houses are visible (a), together with the generated shadows and a part of the road network. In (b), the locations of the identified buildings are shown with their location, size and orientation. Although the phenomenology of the data is very different from the one of industrial buildings for which the algorithm has been developed, the three most evident targets are identified.

- each of the scene objects can be decomposed into a set of elements that get probabilistically mapped into the image data space via a likelihood term that takes into account the geometric and radiometric phenomenologies of sub-metric SAR in urban environments;
- the detection of the buildings is based on a decision likelihood ratio;
- the optimization of the obtained posterior is based on Gibbs sampling and Monte Carlo Markov chains.

The developed algorithm was demonstrated on real data acquired by airborne X band sensors on built-up areas in high complexity settings characterized by tall vegetation and strong metallic scatterers. Further results are presented in chapter III using airborne data with resolutions in the range of 0.5 to 2 meters on a variety of scenes.

Part III

Evaluation and validation of results

Abstract

The novel methods and techniques developed are evaluated, validated and applied to large amounts of data.

The information extraction systems are combined into a complete, automatic data processing chain that serves a knowledge-based information mining and content-based image retrieval system used in the analysis of a nation-wide ERS dataset.

The automatic, space-variant model-order selection capabilities of the extended denoising and information extraction system are demonstrated by generating a local model complexity based classification map from fine-mode RADARSAT data.

The developed system for scene reconstruction based on Bayesian classification and fusion of the extracted features is applied to the analysis of images acquired by the SRTM mission and by the DLR E-SAR system. The scenes are reconstructed by combining the obtained classification and land-use maps with interferometric height data.

Finally, the scene understanding algorithm based on the Gibbs sampling of hierarchical marked point processes is evaluated by applying it to sub-metric Intermap data acquired on an industrial site as well as on high-complexity composed buildings in urban environments.

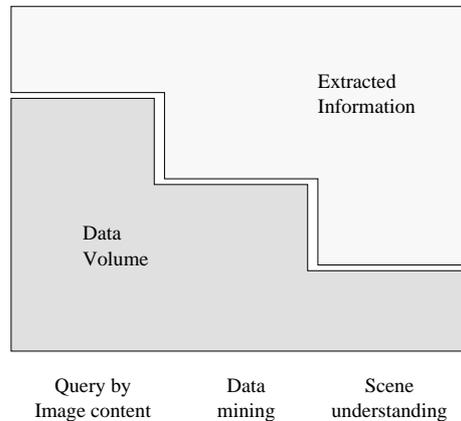


Figure 6.7: The data analysis hierarchy, from content-based query to scene understanding. Due to the problem complexity, the system scales down the amount of data to be analyzed such that the detail of extracted information can be increased.

chapterLarge scale evaluation and validation of developed methods and algorithms **Abstract**

The algorithms and techniques introduced in the last chapter are evaluated and validated on a variety of data including real meter resolution SAR data from airborne sensors serving as a testbed for next generation spaceborne radar systems.

6.10 Content-based retrieval, information mining, scene understanding

Modern remote sensing systems are capable of acquiring and processing very large amounts of data: the SRTM mission collected about 11 TBytes of interferometric SAR data per day. Content-based image retrieval, image information mining and scene understanding all relate to the need of analyzing large amounts of data to produce descriptions of their content in terms of models.

6.10.1 Information extraction for application-independent data characterization

Content based image retrieval, image data mining and scene understanding share common methods (data analysis and probabilistic modeling techniques) and goals (the extraction of meaningful information from data affected by uncertainty), but they operate at different scales of data amount. Content based image retrieval focuses on extracting compact representations for the retrieval of interesting elements in very large amounts of data; image data mining works at an intermediate level, trying to find interesting configurations

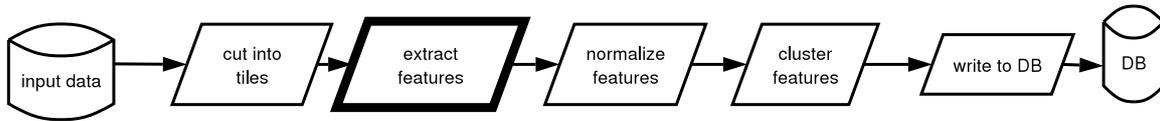


Figure 6.8: Scheme of principle of ingestion chain operations.

in moderately populated feature spaces; scene understanding tries to extract the maximum possible amount of information from a given dataset applying to it models of high complexity.

All three approaches to the analysis of the data, although they operate on different scales of detail and of data amount, share the same need for an application-independent description of the contents of the data. Such description needs to be generated in an automatic way, must be application independent and needs to be as general and complete as possible.

As in the case of the hierarchical model inversion algorithms developed in previous chapters, the analysis is based on a series of description levels linked by conditional probability links. The image data (level 0) is processed to extract primitive features and meta features (levels 1 and 2). Information at level 1 is in form of a parameter vector of a signal model. The specific signal model is the level 2 of information representation. A further description level, a characteristic vocabulary of signal classes (level 3), is generated by unsupervised information clustering for each signal model. User specific interests, that is, cover type labels, are linked to combinations of terms in these vocabularies using simple Bayesian networks. Levels 1 to 3 are obtained in a completely unsupervised and application free way during data ingestion in the system. The information at level 4 can be interactively defined by users employing a learning paradigm. In an additional step of stochastic modeling, the stochastic link between signal classes and user (subjective) labels is described using a vector of hyper-parameters.

6.10.2 The knowledge-based information mining system

The knowledge-based information mining (KIM) system is designed as a system to operate on large archives of remote sensing data in a way that is free from application specificity. After extraction of primitive features takes place as described in previous chapters, the features are grouped by similarity by using a K-means clustering algorithm. The clusters have no direct meaning, since they group points in an n-dimensional space of non-commensurable variables. Still, they represent characteristics of the image seen as a multi-dimensional signal. It is possible to associate meaning with these clusters through training. A user can tell the system that a specific, weighted combination of clusters represents a derived feature of the data. By making this association, it is possible to select all images in the database that have that specific combination and may therefore contain the feature that the user is searching for.

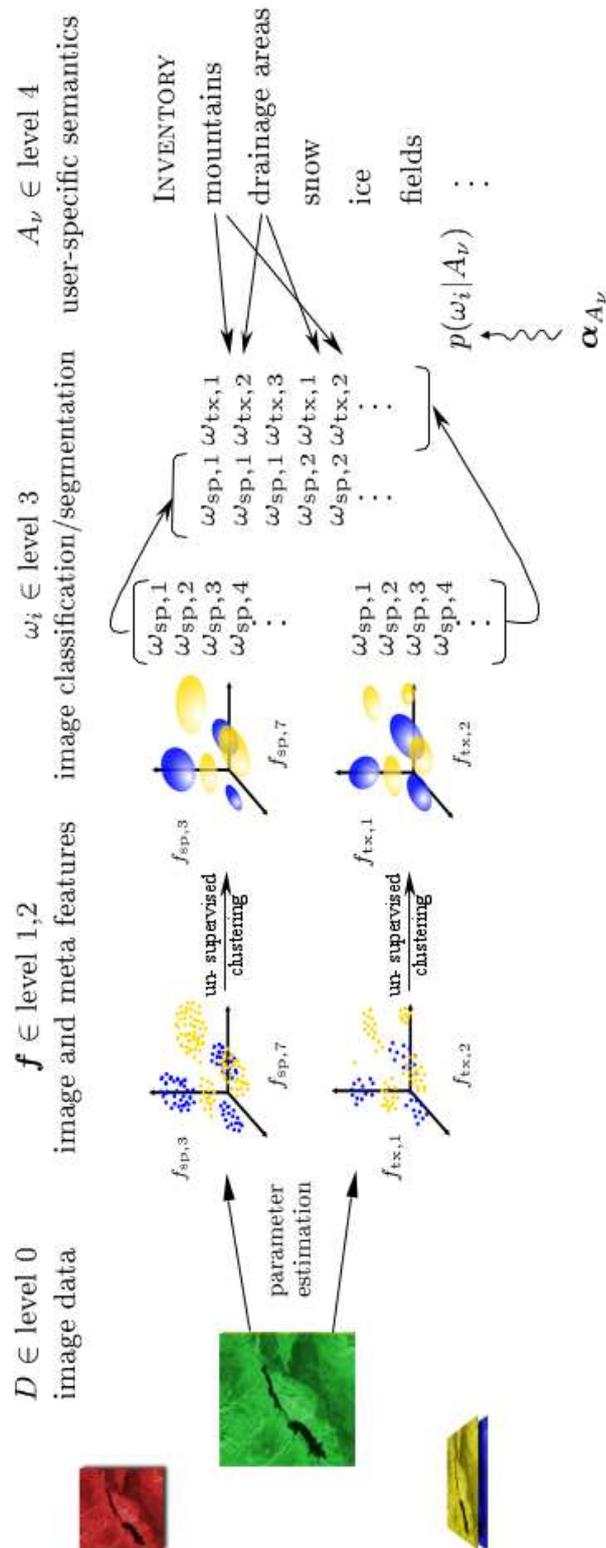


Figure 6.9: Hierarchical modeling of image content and user semantic (Schröder et al., 2000a): description levels are linked by conditional probability links.

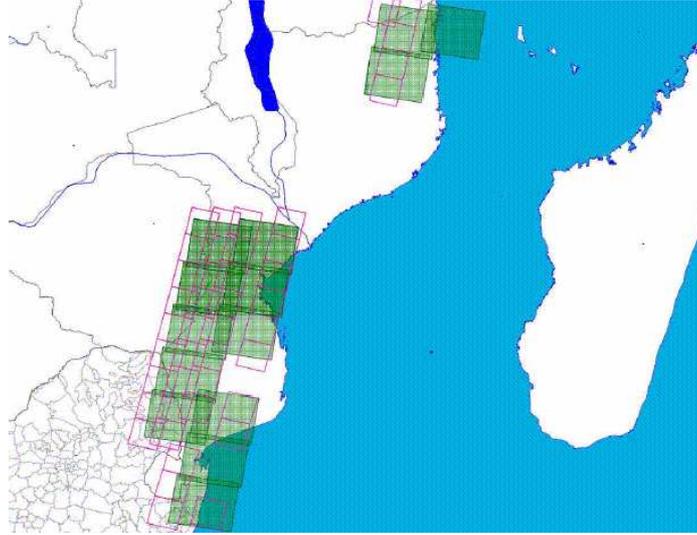


Figure 6.10: KIM coverage of Mozambique.

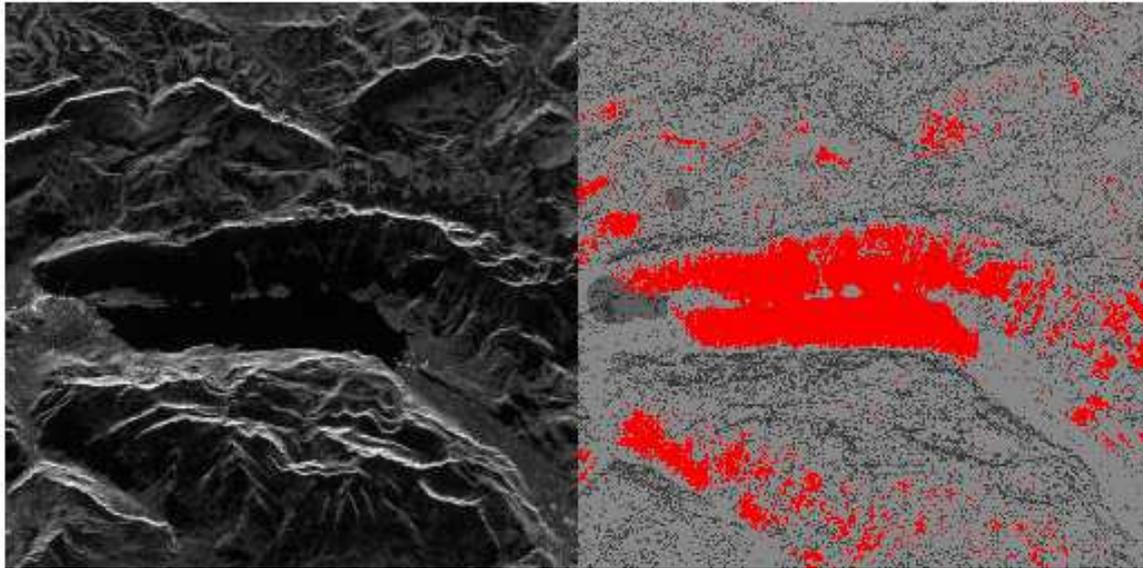
	Kosovo	Mozambique
Sensor	E-SAR	ERS-1
Scenes	1	30
Archive	800 MB	3.5 GB
Mining	24 MB	800 MB
Compression	33	4
Application	scene understanding	information mining

Table 6.1: Summary of SAR data ingested in the KIM system and the level of data analysis considered.

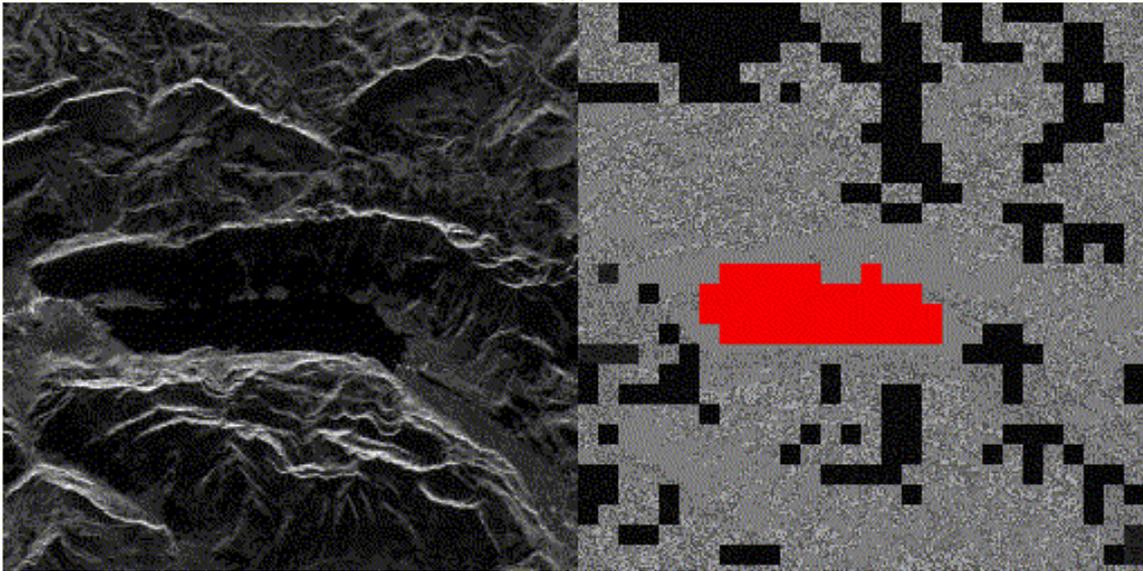
6.10.3 Knowledge-based information mining evaluation results

Table 6.1 presents the summary of all SAR data sets ingested in the KIM system, organized by site. The table outlines the nature of the function implemented in KIM. In the Content Based Image Retrieval mode the compression factor is very high (~ 100), thus enabling the search of very large volumes of data: the accuracy of the search is limited, proportional to the available information for on-line search. The information mining mode is using a moderate compression factor, thus enabling the access to more detailed information for on-line mining. In the scene understanding mode, all image content information is used, thus enabling accurate exploration and interpretation of the images and observed scenes.

The described image information mining based approach to SAR model based despeckling and information extraction evaluation has been applied to large ERS and SIRC/X-SAR coverages of Mozambique and Switzerland. The results obtained allow the navigation of image archives containing 15 *GBytes* of data imaging surfaces that span about $800 \times 600 \text{ km}^2$ and demonstrate the value of the obtained image features.



(a)



(b)

Figure 6.11: (a) Classification based only on backscatter cannot separate lake and mountain shadow in a scene from 75 m resolution X-SAR data. (b) If the fusion of estimated backscatter and texture norm is used, the lake is correctly identified.

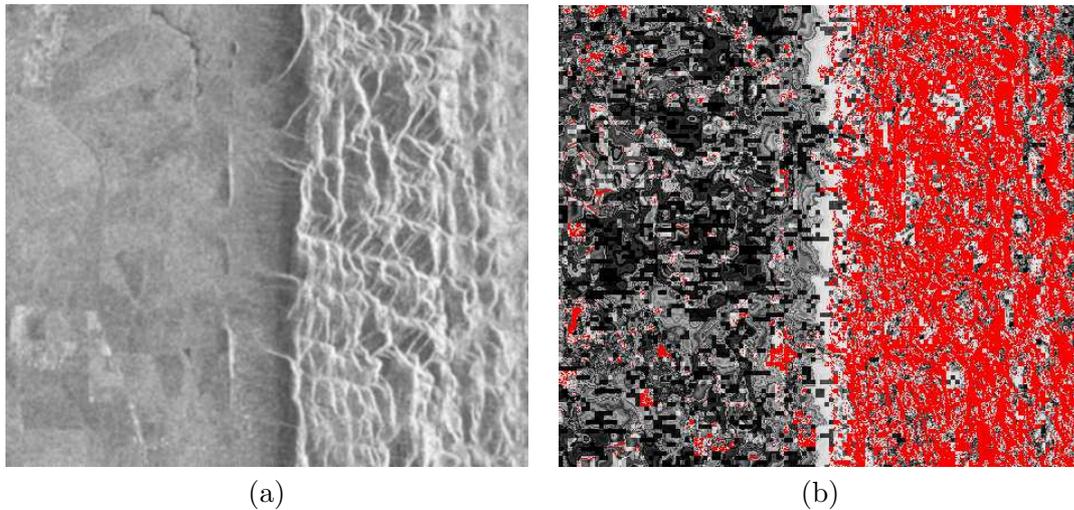


Figure 6.12: (a) SAR relief in an ERS scene on Mozambique at 60 m resolution and (b) corresponding classification show the understanding of geomorphological features using texture estimated at low-resolution scales.

An example of land use cover discrimination is visible in figure 6.11a: even when a classification based only on backscatter cannot separate lake and mountain shadow in a scene from 75 m resolution multilooked X-SAR data, the system is able to integrate such data by fusing the estimated backscatter and the texture norm. By using both information sources, the lake is correctly identified.

In figure 6.12, the SAR relief in an ERS scene on Mozambique at 60 m resolution is reported together with the corresponding classification of geomorphological features using texture estimated at low-resolution scales.

Further results were obtained by analyzing with the pixel-based information extraction algorithm in chapter 4 a meter resolution E-SAR dataset in the frame of the humanitarian demining SMART project (Vanhuysse et al., 2003). The obtained results tend to discriminate in a clear way different types of land cover and scene structures.

The signal-based, level 0 classes considered for information mining and scene understanding from sub-metric resolution polarimetric SAR are all derived from X and L-band SAR polarimetric data. In particular, the despeckled backscattered intensities are separately considered together with a backscatter mean and a backscatter .

The different channels are separately subject to feature extraction and to similarity grouping, and then input into the system and made available for user learning.

Results are shown in figure 6.13 and 6.14. The information made available in the system allows the separation of a number of different cover types as well as the recognition of different scene structures.

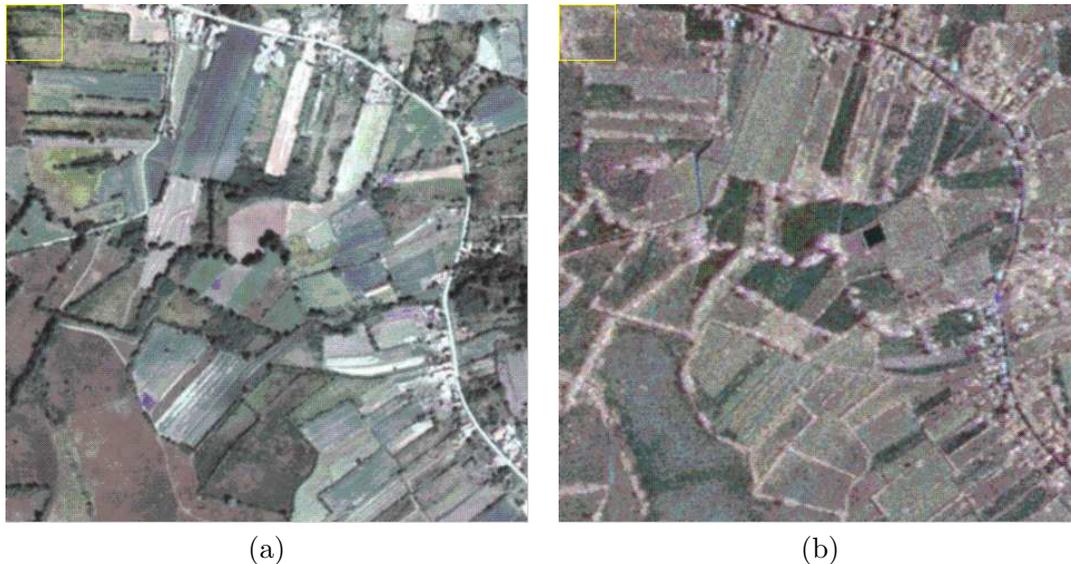


Figure 6.13: Meter resolution information mining example: ground truth Daedalus ATM image (bands 2,3,4) in (a) with Gaussian-normalized RGB quicklook in (b).

6.11 Case studies with local model order selection in model based information extraction

We consider a RADARSAT-1 fine mode four-looks image (in figure 6.15a) of Las Vegas, USA. At least two different kinds of structures with very different target complexities are present in the data. Smooth desert areas lay close to a very complex urbanized environment.

The RADARSAT-1 image in figure 6.15, is filtered by using the model order fixed across the whole image. figure 6.16a is processed assuming a low order GMRF model. This accounts well for the smoothness of the desert area, but tends to smear the borders of the city scene. Figure 6.16b, processed instead with higher model order, tends to provide a better reconstruction of the complex city structures, but produces a sub-optimal filtering of the smooth desert area.

Figure 6.17 shows the automatically estimated local SAR intensity backscatter and the corresponding space variant model order for the Las Vegas RADARSAT-1 scene: the system tends to discriminate the complex urban area from the smoother natural surroundings. Roads and natural relief structures are separated as being more complex than the surrounding desert area. Furthermore, the city area is clearly discriminated from the smoother, simpler surroundings.

6.12 Urban scene reconstruction from InSAR data

A different application of the algorithms presented in previous chapter for the extraction of geometrical information from high resolution SAR and InSAR data is related to the

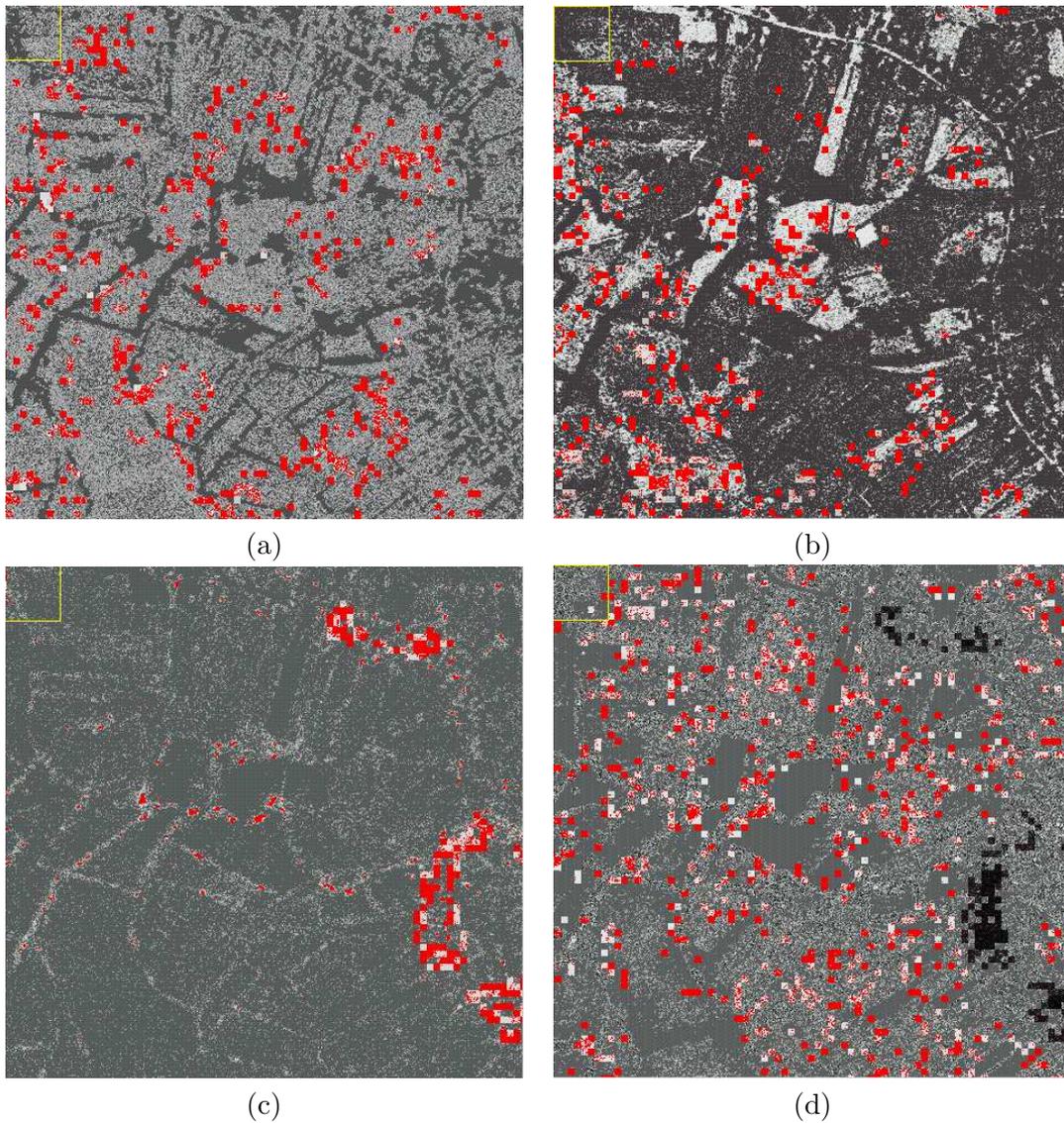


Figure 6.14: Meter resolution information mining example: the original data hh polarized intensity and texture from 6.13b express features that are linked to the user domain by machine learning. In the example, borders (a), fields (b), urban (c), and isolated trees (d).

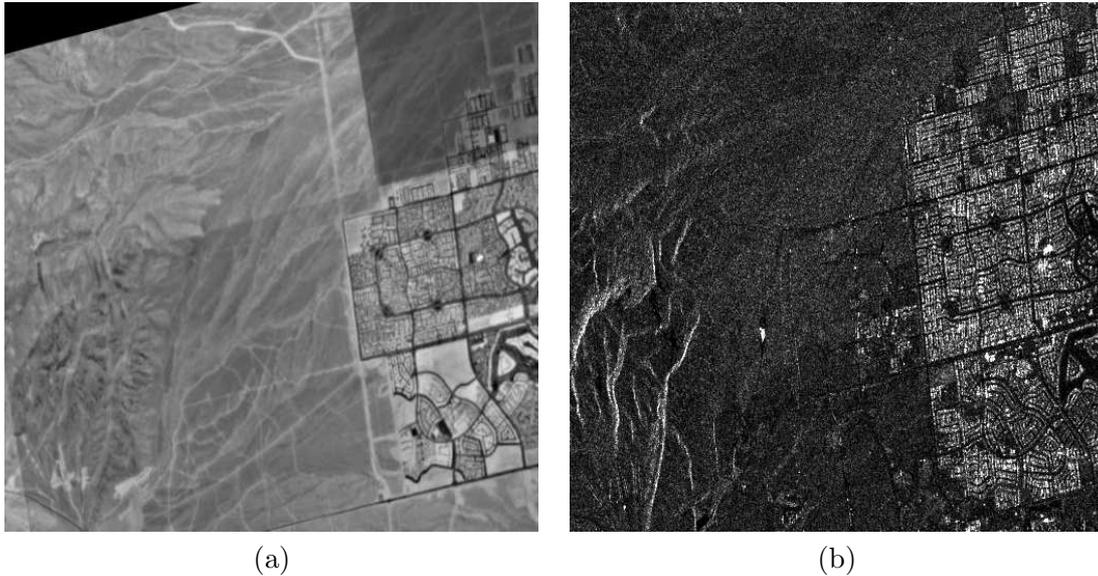


Figure 6.15: Uncalibrated optical (from Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey (Las Vegas aerial photographs, 10 jun 1994)) (a) and RADARSAT-1 fine mode four-looks image (b) of Las Vegas, USA. At least two different kinds of structures with very different target complexities are present in the data. Smooth desert areas lay close to a very complex urbanized environment.

specific scene understanding task of recovering the aspect of an urban scene starting from remote sensing data.

InSAR as a multiple model, high dimensionality data source The scene reconstruction procedure described in chapter 5 is applied to the understanding of urban scenes starting from an interferometric dataset acquired in the context of the Shuttle Radar Topography Mission (SRTM) on the city of Baltimore, USA. The mission, occurred Feb. 11-22, 2000, was based on the single-pass acquisition of multiple streams of data from two sets of antennas on board of the Space Shuttle. The 12 Terabytes of raw data acquired are currently being processed into digital elevation maps, providing the first-ever global coherent DEM of the Earth between latitudes -60° , $+54^\circ$ together with radar backscatter and interferometric coherence information. The dataset under study, with a spatial resolution of about 25 m, was acquired at X band.

A very different dataset, acquired on the outskirts of Munich, Germany, from an Intermap airborne sensor with a spatial resolution of about 0.5 m again at X band, is used as a second example.

We consider the results obtained with this second dataset indicative of the quality of the developed algorithms for their application to the large amounts of data that will be acquired at comparable resolutions by the next generation of satellite-borne SAR systems in the next few years.

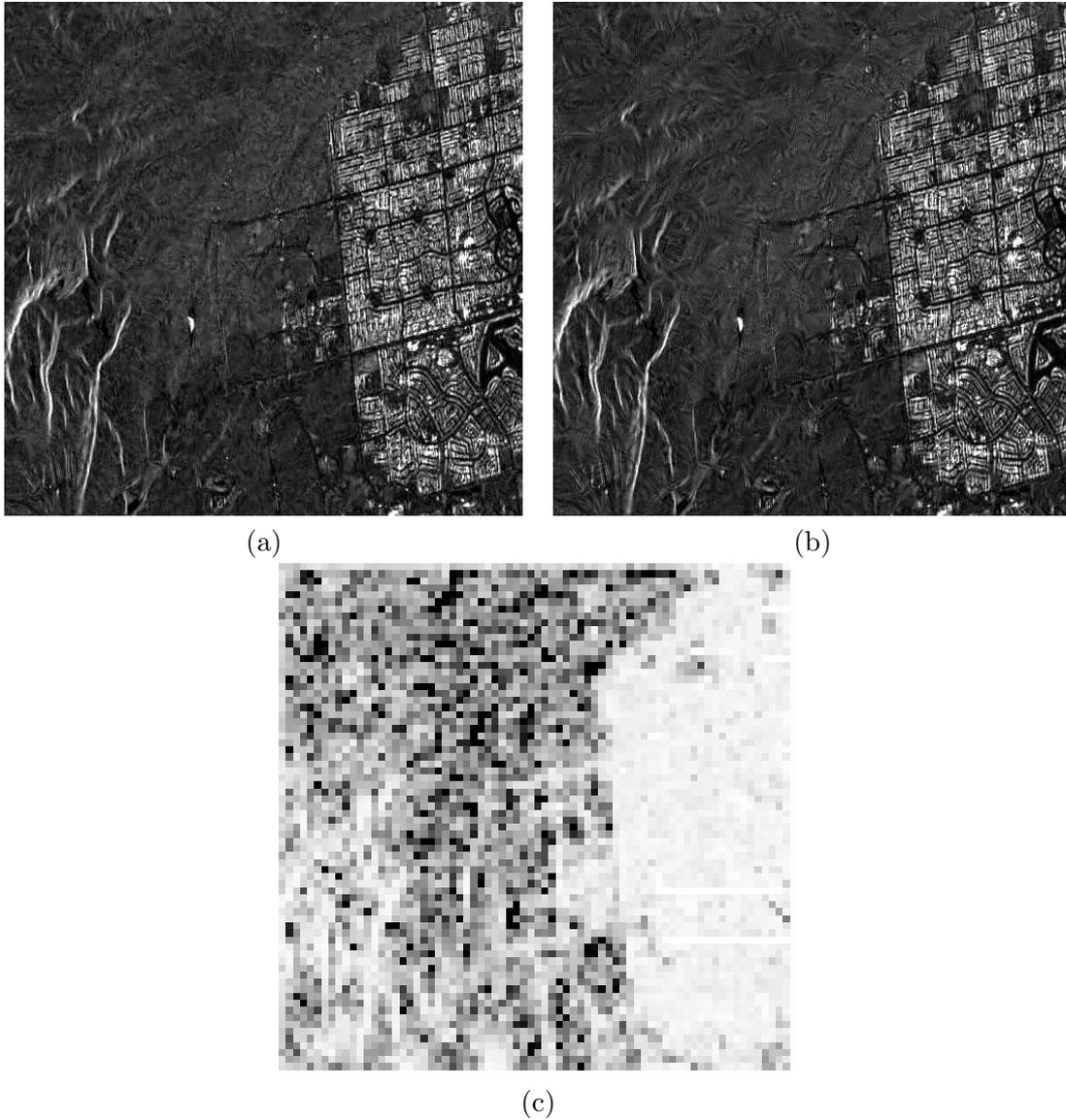


Figure 6.16: The RADARSAT-1 image in figure 6.15, filtered by using the model order fixed across the whole image. The left part of the picture, (a), is processed assuming a low order GMRF model. This accounts well for the smoothness of the desert area, but tends to smear the borders of the city scene. The right part, (b), processed instead with higher model order, tends to provide a better reconstruction of the complex city structures, but produces a sub-optimal filtering of the smooth desert area. In (c), local SNR as estimated number of looks.

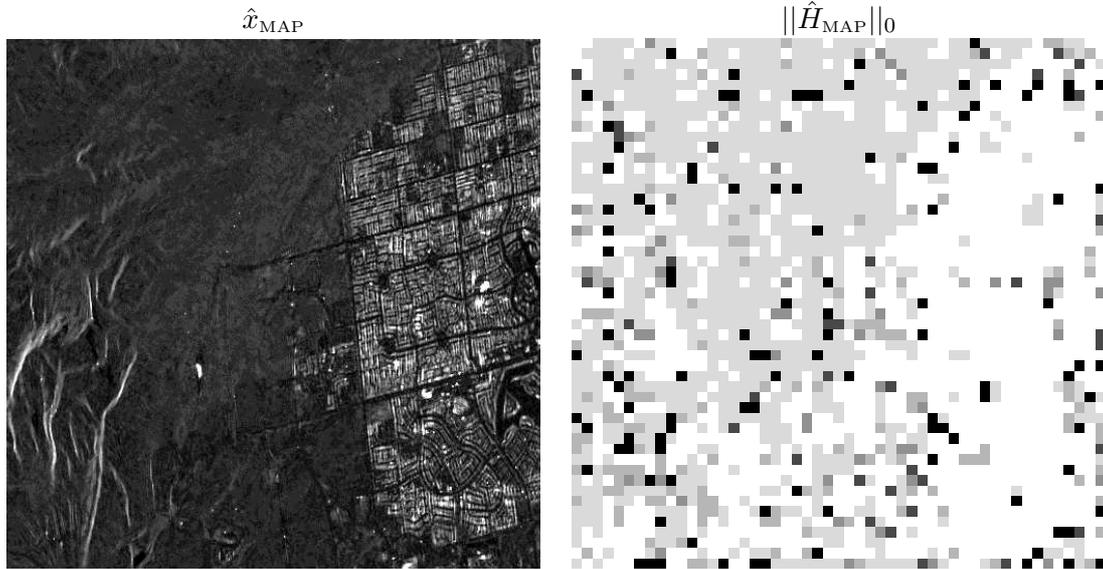


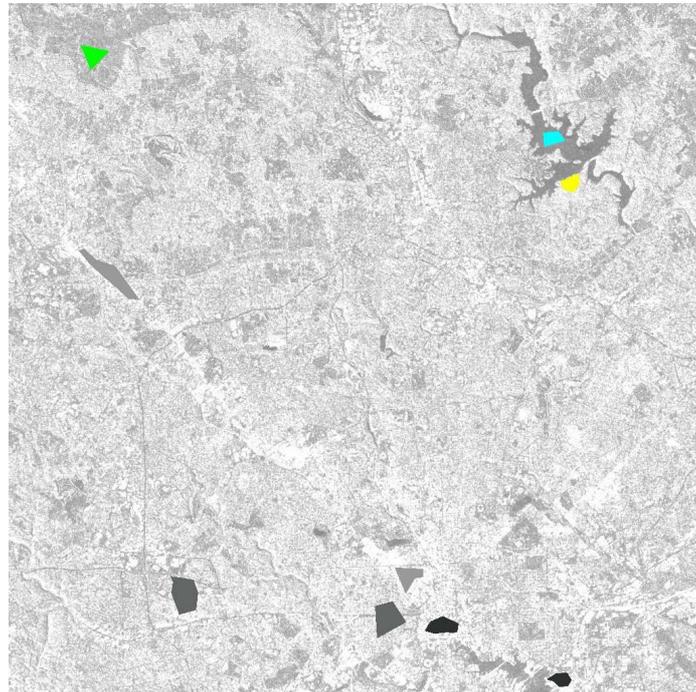
Figure 6.17: Automatically estimated despeckled amplitude and locally estimated maximum evidence model order for the Las Vegas RADARSAT-1 scene: the system tends to discriminate the complex urban area from the smoother natural surroundings. Roads and natural relief structures are separated as being more complex than the surrounding desert area.

6.12.1 Urban land-use mapping from SRTM data

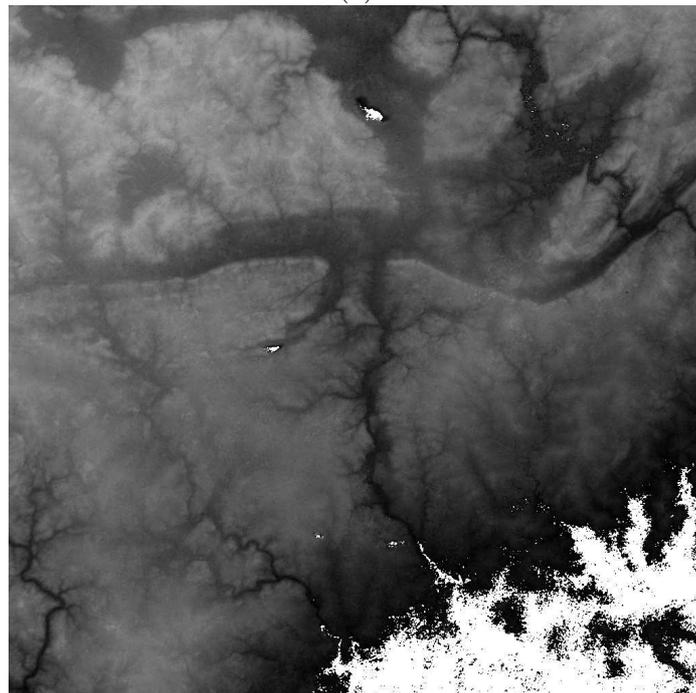
We apply the whole procedure described in chapter 5 — and summarized in figure 5.2 — to urban land use mapping on the city of Baltimore, USA, from an SRTM dataset. A USGS map of the scene is reported in figure 6.20b, while the original backscatter and the DEM data are shown in figure 6.18. No interferometric coherence information is used.

Before being fed into the Bayesian classification and information fusion system, the original dataset layers are subject to separate information extraction and unsupervised classification. A description of the procedure and the intermediate results are shown in figure 6.19.

The information fusion step takes user input in the form of the definition of training regions. The ones defined for the scene under study are reported in figure 6.18a together with their class definitions. The obtained classified results are reported in figure 6.20a together with USGS ground truth: the main directions of development of the city are visible as radial areas of higher building density, while the city center and the areas with highest urbanization around it are well separated from the rest of the urbanized regions. Limited misclassifications are visible in the lower-right part of the dataset where the original height information is corrupted by strong artifacts generated by unmasked water areas. The Loch Raven Reservoir on the upper right part of the image is also recognizable.



(a)



(b)

Figure 6.18: (a) SRTM backscatter intensity image (resolution 25 m, number of looks 4) on the city of Baltimore, USA. A strong multiplicative noise called speckle generated from the coherent nature of the SAR system is visible. Supervised classification/feature fusion training is superimposed on the data: FOREST (green), WATER (light blue), SPARSE_URBANIZED (middle grey), MIDDLE_URBANIZED (light gray), DENSE_URBANIZED (dark grey), AGRICULTURAL (yellow). (b) corresponding SRTM height image on the city of Baltimore. Artifacts on unmasked water areas are visible.

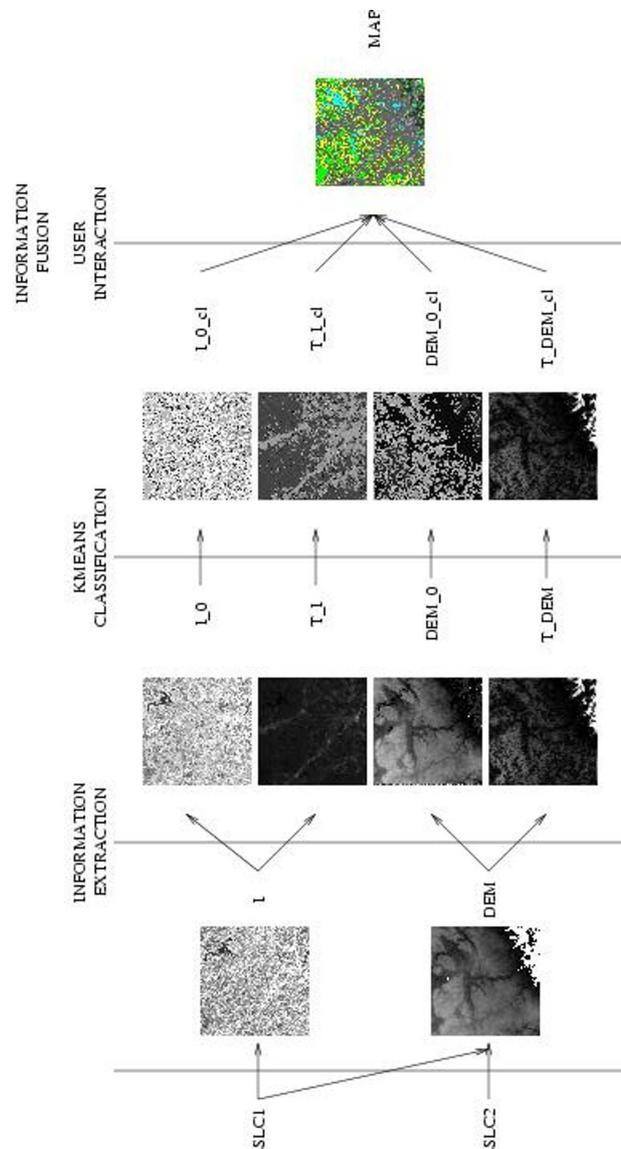
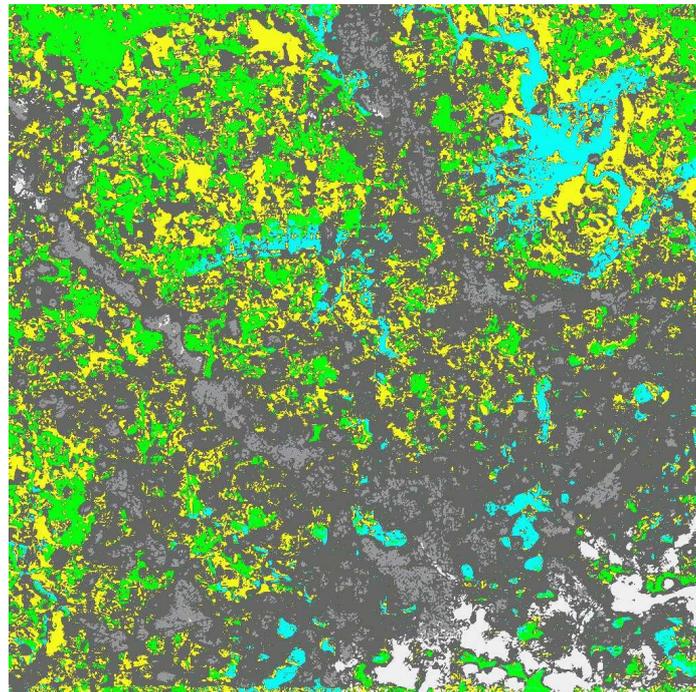
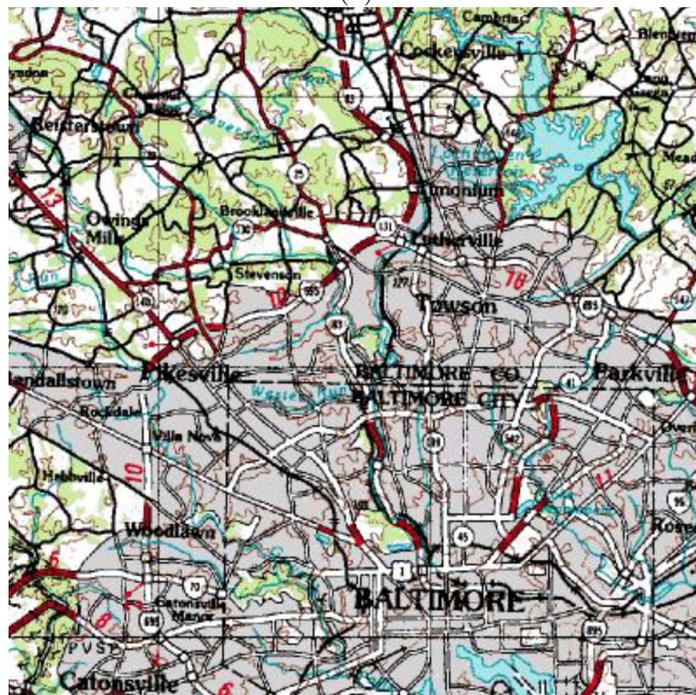


Figure 6.19: The scene reconstruction procedure for urban land use mapping on the Baltimore scene with intermediate and final processing results: the original backscatter intensity and elevation data are processed into their clean estimates and textural descriptions through the information extraction system. The features obtained are grouped by similarity and fed into the Bayesian classification/fusion system together with the training in figure 6.18a to obtain the final land use map.



(a)



(b)

Figure 6.20: (a) Bayesian classification/fusion results: FOREST (green), WATER (light blue), SPARSE_URBANIZED (middle grey), MIDDLE_URBANIZED (light grey), DENSE_URBANIZED (very light grey), AGRICULTURAL (yellow). (b) USGS 01 Jul 1984 ground truth (from Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey (Baltimore map 01 jul 1982)). The main directions of development of the city are visible as radial areas of higher building density, while the city center and the areas with highest urbanization around it are separated from the rest of the urbanized regions. Misclassifications appear on areas where the DEM information is corrupted by artifacts.

6.12.2 Large building recognition from SRTM data

Although the limited spatial resolution (25 m) of the SRTM dataset limits its applicability to the understanding of very complex urban environments, we evaluate the results of the application of the described scene reconstruction framework to the identification of large buildings.

The area of the city that we concentrate our analysis on (about 10x10 km wide) is reported in figure 6.22b: it includes a city park as well as highly developed urban elements. Original backscatter intensity and heights are reported in figure 6.21.

The input training regions provided to the interactive Bayesian information fusion system are described in figure 6.21a. The results of the procedure are reported in figure 6.22: the different elements in the scene are discriminated from each other, and many individual buildings are separated.

A perspective view of center portion (about 2.5 km of side) of the local DEM where the heights of class BUILDING have been exaggerated by 10% (in figure 6.23c) shows the distribution and the 3D aspect of some of the streets in the scene.

A comparison with the ground truth in figure 6.23b shows that all but one of the buildings reported in the 1986 USGS ground truth map are recognized as such. The remaining detected buildings, many of which are aligned to the existing road network, are either actual buildings erected after the ground truth map was produced or false alarms.

6.12.3 Building recognition from Intermap data

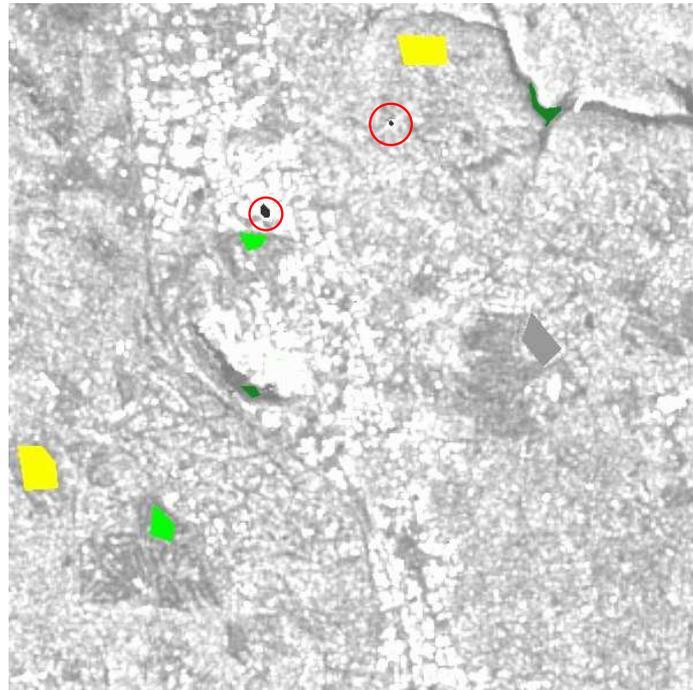
While the resolution of the SRTM system makes it more appropriate for the study of large scale natural structures than for the understanding of complex urban environments, the next generation of meter-resolution spaceborne SAR sensors will be producing on a regular basis vast amounts of detailed data on urban environments.

Despite differences in some aspects of the sensor characteristics, current airborne systems provide a test bed for the evaluation of scene reconstruction algorithms to next-generation spaceborne data.

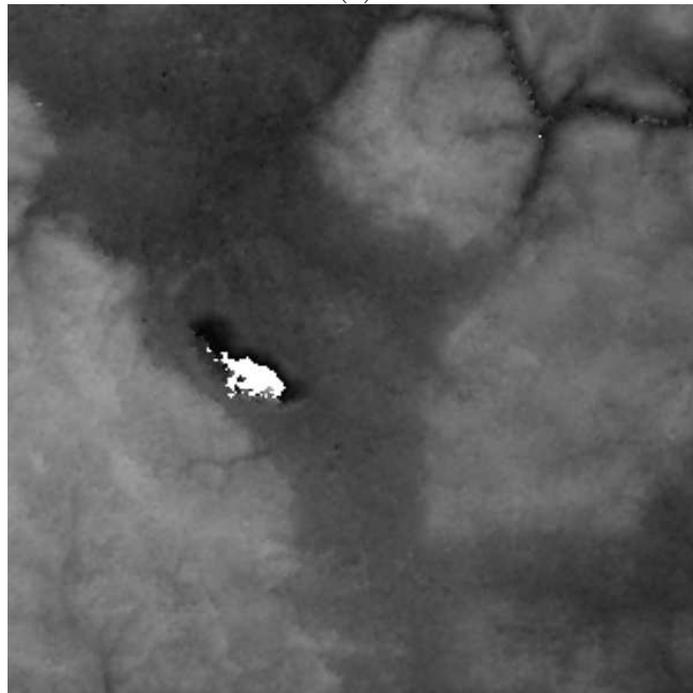
An Intermap X band interferometric dataset with resolution of about 0.5 m taken during a flight on a rural area east of the city of Munich, Germany, is used in a further building recognition experiment.

The test scene has very high complexity. The structures to be reconstructed are small houses partially hidden at X band by surrounding taller vegetation (see figure 6.24). Typically for very high-resolution SAR data in built-up environments, multiple scattering and other disturbing effects tend to dominate the data, saturating the receiver and propagating to different areas in the image (figure 6.25).

The signal features used include the interferometric coherence, interferometric phase gradients obtained via Gabor filters and single image intensities and intensity textures as well as interferometry-derived measures of the local elevations. The signal features are

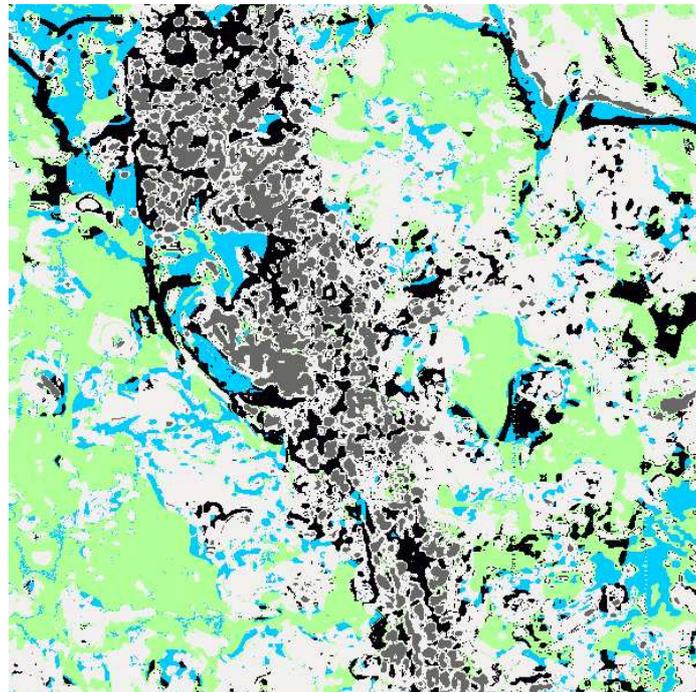


(a)

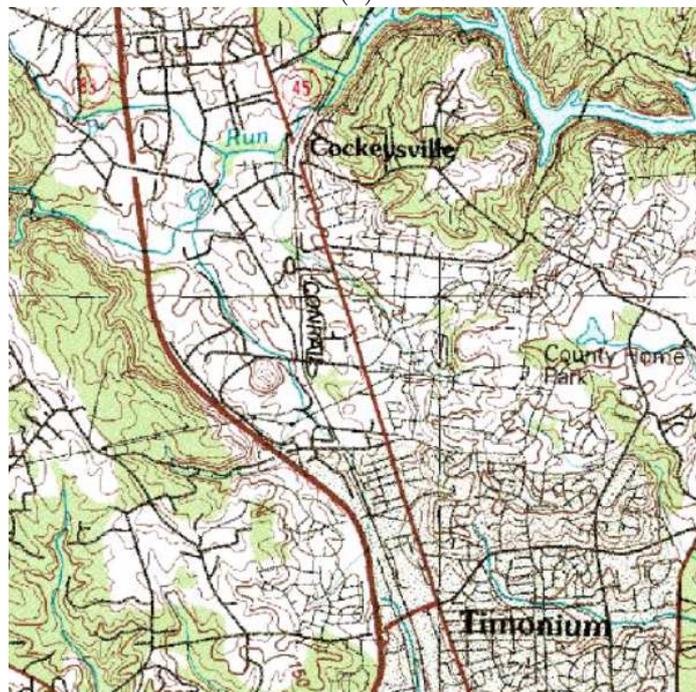


(b)

Figure 6.21: (a) SRTM backscatter intensity image on the area object of building recognition with superimposed Bayesian classification/fusion training: BARE-AGRICULTURAL (yellow), BUILDING (dark grey in red circles), WATER/SHADOW (dark green), GENERAL_URBANIZED (light grey), FOREST (light green). (b) SRTM height image on the same area. Artifacts are visible on unmasked shadow/water areas.



(a)



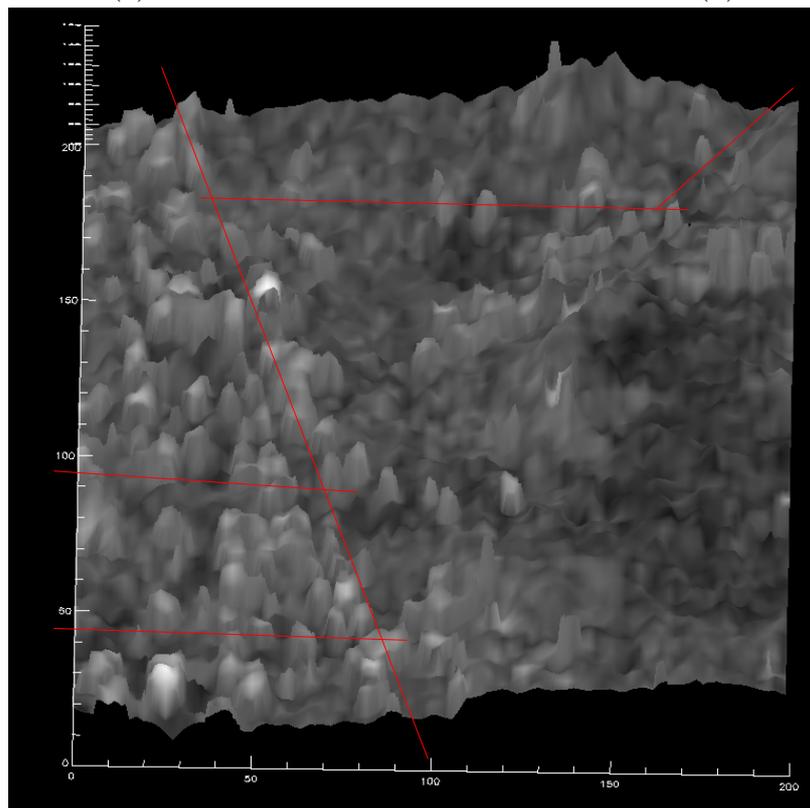
(b)

Figure 6.22: (a) Bayesian classification/fusion results: BARE-AGRICULTURAL (light green), BUILDING (grey), WATER/SHADOW (light blue), SHADOW (black), GENERAL_URBANIZED (white), FOREST (green) and (b) USGS 01 Jul 1984 ground truth (from Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey (Baltimore map 01 jul 1982)).



(a)

(b)



(c)

Figure 6.23: (a) Bayesian classification/fusion results with in transparency (b) USGS 01 Jul 1986 ground truth (from Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey (Baltimore map 01 jul 1982)). All but one of the buildings reported on the 1986 map are recognized as such. The remaining detected buildings are either actual buildings erected after the ground truth map was produced or false alarms. (c) 3D view of center portion (about 2.5km of side) of the local DEM where the heights of class BUILDING have been exaggerated by 10% with superimposed annotation data. The distribution of some of the streets in the scene is visible.



Figure 6.24: Views of the rural test area: buildings are partially hidden by trees at X band. Viewpoints are indicated in figure 6.25b.

separately subject to feature extraction and unsupervised classification and are subsequently fused in a supervised way after specifying the training mask in figure 6.25a.

Although the scene is very complex and the user provided training is intentionally limited to two very simple areas describing respectively built-up and natural scene elements, the results tend to separate built-up elements from different image objects such as trees or vegetation. Buildings in the scene are marked as such, while some spurious classification in the middle of the scene is determined by responses by greenhouses and other non-permanent structures. The results obtained are shown together with ground truth in figure 6.26.

6.13 Marked point process model–based building reconstruction from metric airborne SAR

The results obtained with pixel–based methodologies in figure 6.26 are of limited interest because they fail to decompose the scene into a set of different, separated scene objects. We apply therefore the object–based methodology introduced in chapter 6 to the analysis of very high resolution SAR data.

6.13.1 Industrial scene understanding from metric SAR backscatter

A single intensity image with a spatial resolution of 2 meters acquired on the German Oberpfaffenhofen site by the Intermap airborne X band SAR system is used for the validation of the proposed SU strategy.

The scene contains a number of large parallelepipedal industrial buildings surrounded by elements of the transport infrastructure.

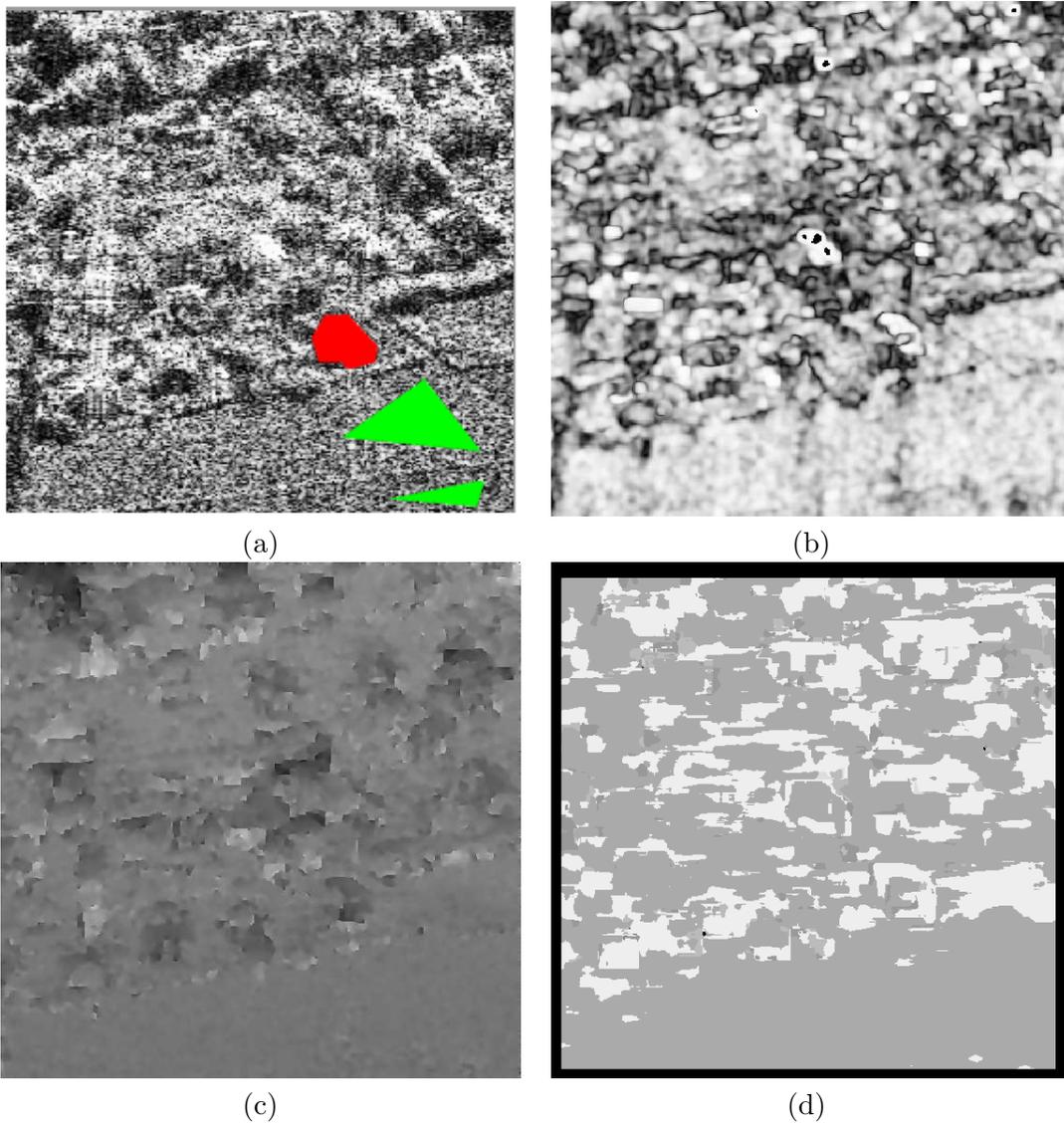


Figure 6.25: Rural test scene data: (a) intensity image of the site presented in figure 6.24. The intentionally very limited training areas are chosen to allow the detection of buildings by separating them from other scene structures. The information on buildings is extracted from intensity image, coherence (b), interferometric-derived elevations (c) and interferometric phase gradients (d), and further fused according to the “example” given by the training area. The cross-track direction is indicated as an aid in the interpretation of the data, given the strong dependence of meter-resolution SAR phenomenology on the acquisition geometry.

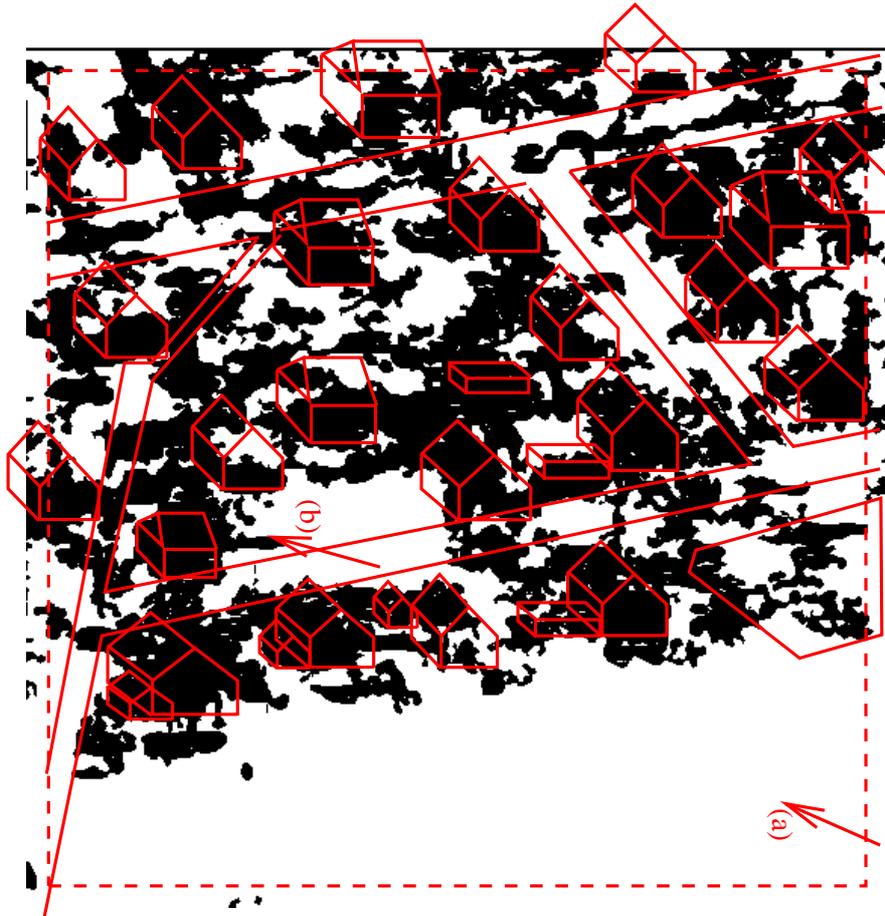


Figure 6.26: Rural test scene results: (a) extracted building map with ground truth provided for comparison. The arrows mark the view points for the photographs in figure 6.24 while the red box marks invalid results on image borders.

The data show well visible artifacts related to the high resolution of the sensor used and to the smooth, reflective nature of the buildings in the scene. In particular, radiometric saturation extending on many image areas by very strong dihedral scatterers affects most part of the data.

The object model considered constrains it to be a parallelepipedal building that gets mapped in the data into a direct edge and a dihedral scattering facet only.

The original image and the obtained results are reported in figure 6.27.

The main buildings in the data are identified. While their geometry is generally consistent with the scene, the dimensions in the range direction are exaggerated due to the influence of SAR phenomenology: strong scatterer responses tend to appear elongated in the range direction, acting as features that suggest an effective building shape that is wider in that direction. There are clear errors in the upper building on the left, where the reconstruction does not seem well related to the original image.

6.13.2 Composed building understanding from sub-metric SAR backscatter

A similar dataset with a resolution of 0.5 meters acquired on the Dutch city of Maastricht by the Intermap airborne sensor is used as a second example.

The scene is much more complex than in the case of the first dataset. It contains a museum building situated near to the river Maas crossing the city, visible in the lower right corner of the image. The areas surrounding the building are covered by low vegetation. The building is composed of a strongly reflective, round metallic tower with uneven surfaces. The main body of the building consists of a closely spaced array of tall, thin block buildings that tend to occlude each other when observed through the sensor. Even with ground-truth optical data available, the analysis of the amplitude SAR data is complicated by a number of artifacts that are typical of sub-metric resolution SAR. In particular, radiometric saturation effects, layover from the main building body and multiple reflections are clearly visible. A separated set of multiple returns originates from clear edges, early direct echo returns and dihedral scattering on the base of the building are observed as well. The models used for the scene have to account for the complexity of the building. In particular, parallelepipedal and round tower objects are considered together with models for houses with a V-shaped roof.

The results are reported in figure 6.28.

The limitations of the method are both evident: although the principal elements in the scene are correctly identified, the reconstructed main building body is insufficient because of the fact that the actual structure of the building is not observable in the radar return. Furthermore, local along-range radiometric saturation completely hinders the backscattering from one of the edges of the object. In other areas, the return from the far-range edge of the building roofs might be invisible: the lack of an indication of the extent of the object has a direct effect on the validity of the results just as it would constitute a limit to human interpretation of the data. It is clear that since the presented algorithm is designed as a tool

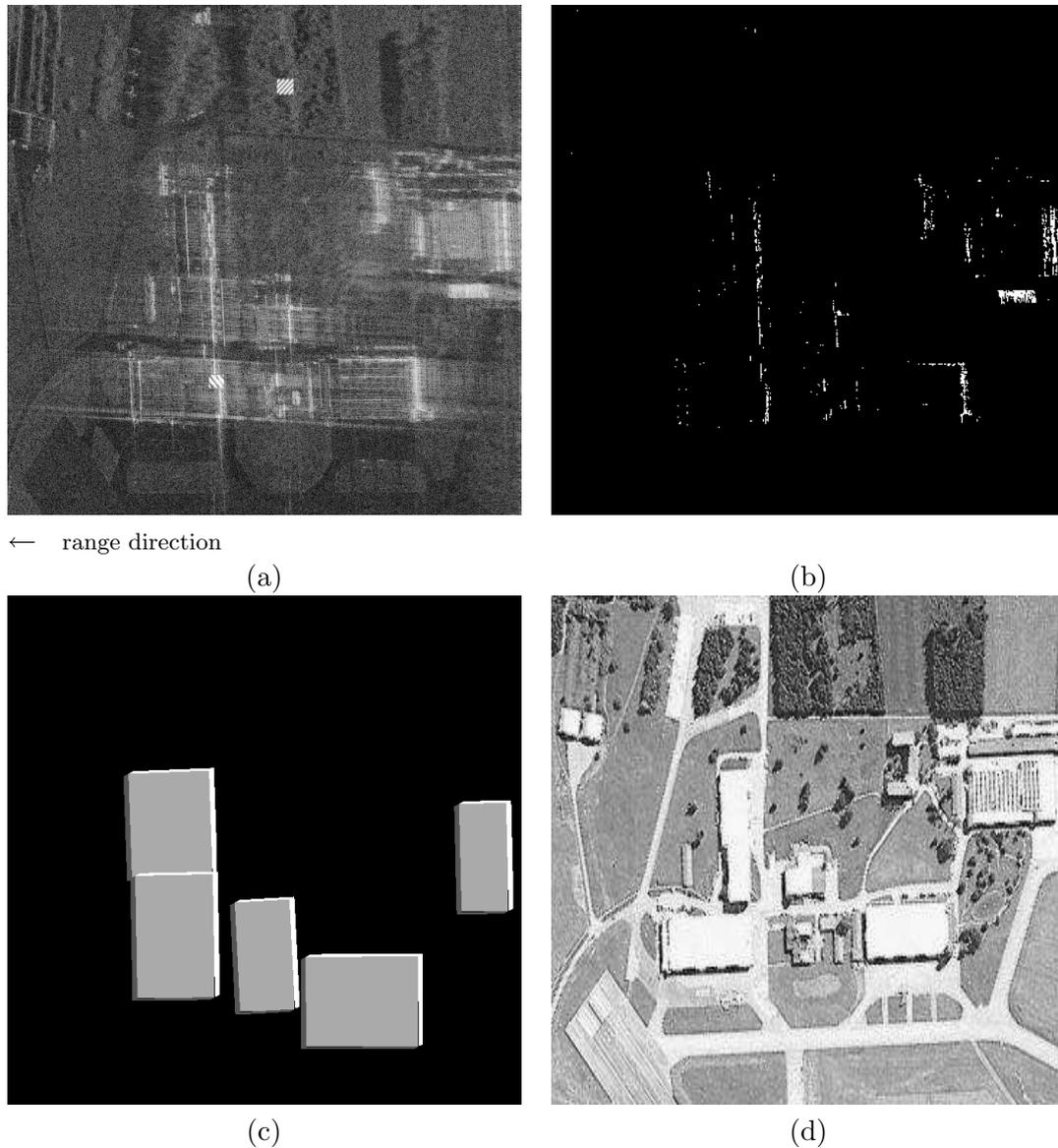


Figure 6.27: (a) Input 8 looks intensity SAR data on the Oberpfaffenhofen, Germany site with spatial resolution of 2 meters. Growing range in right-to-left direction. Training areas for clutter, building response are marked by short oblique white lines. (b) likelihood ratio classification for the data (c) results on the Oberpfaffenhofen area: most buildings are identified, while radiometric distortion in the range direction explains the missed elements. The image represents the reconstructed scene model with each of the objects represented by a set of facets plus a front and a shadow region in an optical-like view; (d) aerial image of the relevant area, in ground range.

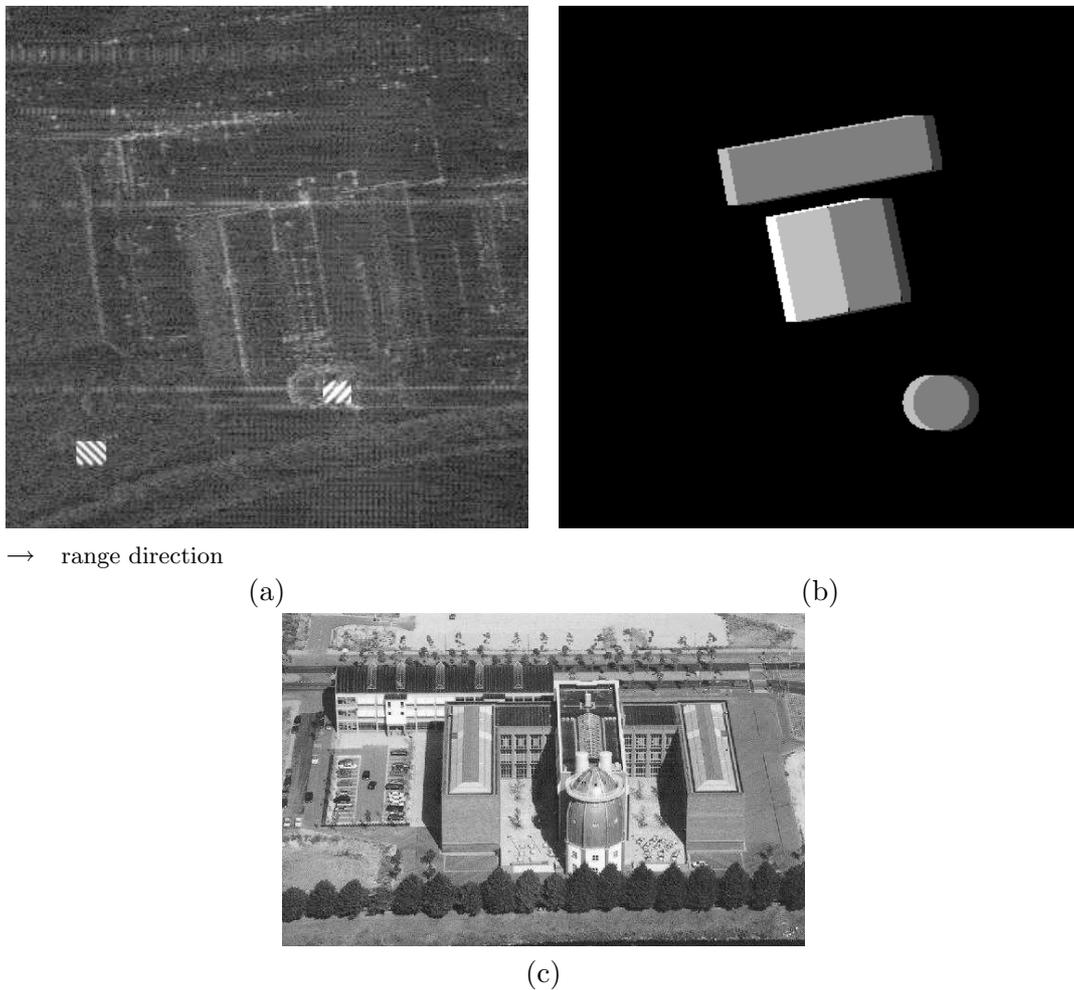


Figure 6.28: (a) Input intensity SAR data on the Maastricht site with spatial resolution 0.5 meters Growing range in left-to-right direction. Training areas for clutter, building response are marked by short oblique white lines. (b) results on the Maastricht area. The image represents the reconstructed scene model with each of the objects represented by a set of facets plus a front and a shadow region in an optical-like view. Along-range radiometric distortion due to saturation completely hinders the backscattering from one of the edges of the main building object, producing an incorrect reconstruction of the close array of box-shaped buildings that tend to occlude each other when observed through the sensor; (c) ground truth.

for the investigation of the limits and merits of information extraction from a single high resolution intensity image rather than for building reconstruction, application scenarios will need the integration of its output with data derived from different sources.

6.14 Summary

In this chapter, the following points have been considered:

- the algorithms and techniques introduced in the last chapters were evaluated and validated on real high resolution SAR data from airborne sensors as a testbed for next generation spaceborne radar;
- a number of approaches were taken in the validation, making use of large quantities of data. In particular, an approach based on the evaluation of the quality of the different extracted image descriptors was introduced that makes use of image information mining techniques: applications related to the handling of large archives in Content-based image queries as well as to the detailed analysis of single scenes were considered;
- furthermore, the geometrical features extracted by the data were considered as a basis for the further reconstruction of a number of urban scenes starting from InSAR and SAR data at resolutions ranging from 25 meters to 0.5 meters.

Conclusions

In this thesis, structured hierarchical Bayesian models and estimators were considered for the analysis of multidimensional datasets representing high complexity phenomena.

The analysis was motivated by the problem of urban scene reconstruction and understanding from meter resolution InSAR data, focussing on the inversion of data acquired by the observation of highly diverse, structured settlements through sophisticated, coherent radar based instruments from airborne or spaceborne platforms at distances of up to hundreds of kilometers from the scene.

Based on a Bayesian analysis framework, stochastic models were developed for both the original signals to be recovered (the original scene characteristics that are object of the analysis— 3D geometry, radiometry in terms of cover type) and the noisy acquisition instrument (a meter resolution SAR interferometer). The models were then combined to provide a consistent description of the acquisition process. to be inverted by the application of the so called Bayes' equation.

The developed models for both the scene and the acquisition system were splitted into a series of separated layers with likelihoods providing a probabilistic link between the layers and with Maximum A Posteriori Bayesian inference as a basis for the estimation algorithms. To discriminate between different prior scene models and provide the necessary ability to choose the most probable model for the data in a given set, a Bayesian model selection framework was considered. This approach resulted in a set of algorithms for the analysis of meter resolution SAR and InSAR data acquired on urban environments.

In particular, a set of existing Gauss–Markov random field model–based algorithms for SAR and InSAR information extraction and denoising were extended by space–variant automated model–order selection capabilities whose performance was demonstrated by generating and validating model–complexity based classification maps of a set of test images as well as on real SAR data.

Based on that, a method for building recognition and reconstruction from InSAR data centered on Bayesian information extraction and data classification and fusion was developed. The system integrates signal based classes and user conjectures, and is demonstrated on input data ranging from on board Shuttle based observations of large urban centers to airborne data acquired at sub–metric resolutions on small rural centers.

To overcome the limitations of pixel based models and inference methods, a system based on stochastic geometry, decomposable object Gibbs fields and Monte Carlo Markov chains

was developed and validated on sub-metric data acquired on both urban and industrial sites.

The developed algorithms were then extensively validated by integrating them in an image information mining system that enables the navigation and exploitation of large image archives based on a generic probabilistic characterization of the data that is performed automatically.

Summary of the obtained results

First, the properties of SAR and InSAR systems at metric resolutions in urban environments were introduced as a motivation for the employment of Bayesian analysis for high complexity hierarchical models. The processing techniques that are used to treat the very peculiar geometric and radiometric phenomenology of SAR and InSAR systems were introduced. The system properties were analyzed at metric resolutions in urban environments, investigating geometric effects such as layover, shadowing and occlusion, the radiometric effect of smooth surfaces on speckle noise, backscatter texture and signal-to-noise ratio, the appearance of multiple reflections of the incident radar beam, the various classes of strong isolated scatterers and the peculiar statistics of meter resolution data. Traditional, simulation- and model-based inversion algorithms in the literature were finally evaluated and compared.

Bayesian modelling and estimation techniques for the analysis of multidimensional fields were introduced. The properties of Gibbs-Markov fields and hierarchical Bayesian models were detailed. Bayesian model selection was presented together with the role of the Occam factor and of evidence maximization in model choice and analysis. Bayesian estimation and decision theories were introduced focussing on Maximum A Posteriori estimates and on decisions as posterior ratios. Posterior optimization techniques based on complete enumeration, on Iterated Conditional Modes and on the Gibbs Sampler and Monte Carlo Markov chains were then introduced and detailed.

In the second part, dedicated to novel methods developed for the hierarchical modelling and the inversion of SAR and InSAR datasets, an extension of the image denoising and information extraction method of Walessa and Datcu (2000) and Datcu et al. (1998) was introduced: it incorporates a further estimation level that allows the space-variant selection of the model-order by means of the maximization of the Bayesian model evidence. This allows the system to properly model and reconstruct images that are composed of instances of different stochastic image models with varying degrees of complexity. The extended algorithm was demonstrated on simulated data generated with well-known model orders as well as on Brodatz textures corrupted by multiplicative speckle noise. We gave examples of classification by model order in the context of Gauss-Markov random field model based image denoising and information extraction starting from specked Synthetic Aperture Radar data: classification maps based on model order were therefore generated that demonstrate the discrimination potential of the Bayesian evidence maximization with respect to model choice. The extended system is able to extract descriptors from the data in terms of a

model-based vocabulary, to choose models for the data in a space-variant way in a given reference class and to discriminate different image elements by their complexity, again in terms of the model.

Based on this, a novel framework for scene understanding from InSAR data was presented that is based on Bayesian machine learning and information extraction and fusion in order to complement the strengths and limitations of different data sources. The system is based on a hierarchical model of the acquisition process that leads from the scene to the data: a generic description of the data in terms of multiple models is automatically generated from the original signals, based on modifications of the denoising and information extraction algorithms presented in chapter 4. The obtained feature space is then mapped by the application of simple Bayesian networks to user semantics representing urban scene elements in a supervised step. The developed system was demonstrated by generating high-resolution land-use maps on urban environments from real InSAR data acquired from spaceborne sensors.

Subsequently, to overcome the limitations of lattice-based models, an algorithm for the reconstruction of building structures from a single High Resolution SAR intensity image was introduced, described and demonstrated on data with resolutions ranging from 0.5 to 2 meters. The algorithm is based on the Maximum A Posteriori estimation via Monte Carlo methods of an optimal scene described in a hierarchical Bayesian framework as a set of pairwise interacting Poisson distributed marked points that represent parametrically described scene buildings. Each of the scene objects can be decomposed into a set of elements that get probabilistically mapped into the image data space via a likelihood term that takes into account the geometric and radiometric phenomenologies of sub-metric SAR in urban environments. The detection of the buildings is based on a decision likelihood ratio. The optimization of the obtained posterior is based on Gibbs sampling and Monte Carlo Markov chains. The developed algorithm was demonstrated on real data acquired by airborne X band sensors on built-up areas in high complexity settings characterized by tall vegetation and strong metallic scatterers. Further results are presented in chapter III using airborne data with resolutions in the range of 0.5 to 2 meters on a variety of scenes.

Evaluation of the obtained results and outlook

The algorithms and techniques introduced were evaluated and validated on real high resolution SAR data from airborne sensors as a testbed for next generation spaceborne radar. Large quantities of data and a number of data analysis approaches were taken in the validation. In particular, an approach based on the evaluation of the quality of the different extracted image descriptors was introduced that makes use of image information mining techniques: applications related to the handling of large archives in content-based image queries as well as to the detailed analysis of single scenes were considered.

Appendix A

The K-means clustering algorithm

One of the basic problems that arise in a great variety of fields, including pattern recognition, machine learning and statistics, is the so-called clustering problem. The fundamental data clustering problem may be defined as discovering groups in a data or grouping similar objects together. Each of these groups is called a cluster, a region in which the density of objects is locally higher than in other regions.

The K-means algorithm (Marroquin and Girosi, 1993; MacKay, 2002) partitions the points in the data matrix X into k clusters. This iterative partitioning minimizes the sum, over all clusters, of the within-cluster sums of point-to-cluster-centroid distances. By default, it uses squared Euclidean distances.

Although this algorithm is known for its robustness, it is widely reported in literature that its performance depends upon two key points: initial clustering and instance order.

After the k means have been initialized (Pen et al., 1999), the K-means is an iterative two-step algorithm. In the assignment step, each data point n is assigned to the nearest mean. In the update step, the means are adjusted to match the sample means of the data points that they are responsible for. The algorithm converges when point assignments are unchanged so the means remain unmoved when updated. The K-means algorithm always converges to a fixed point.

- | |
|---|
| <ul style="list-style-type: none">• initialize solution;• loop until termination condition is met:<ul style="list-style-type: none">– for each pixel in the image, assign that pixel to a class such that the distance from this pixel to the center of that class is minimized;– for each class, recalculate the means of the class based on the pixels that belong to that class. |
|---|

Table A.1: A simple version of the K-means clustering algorithm.

Appendix B

Non-analytical optimization and MAP estimation by simulated annealing

The simulated annealing optimization algorithm exploits an analogy between the way in which a metal cools and freezes into a minimum energy crystalline structure (the annealing process) and the search for a minimum in a more general system.

The algorithm is based upon that of Metropolis et al. (1953), which was originally proposed as a means of finding the equilibrium configuration of a collection of atoms at a given temperature. The connection between this algorithm and mathematical minimization was first noted by Pincus (1970), but it was Kirkpatrick et al. (1983) who proposed that it form the basis of an optimization technique for combinatorial problems.

A simulated annealing procedure can be adopted for the estimation of the Maximum A Posteriori: the multivariate posterior is maximized by iteratively evaluating many possible solutions in a multidimensional parameter space. In stages, the optimization algorithm restricts the search paths to only the most promising solutions. While at the beginning of the iterative optimization procedure the sampled posterior function is essentially flat, at the end of the annealing the algorithm is forced to follow a steepest ascent trajectory. The net effect is that of maximizing the $f(\cdot)^{1/T}$ for T that slowly changes from high values to zero to be able to find the maximum of the function.

An acceptance ratio that is dependent on the rate of improvement in the solution is used together with a decreasing “temperature” acceptance factor to probabilistically accept or reject new solutions.

A simulated annealing program consists of a pair of nested DO-loops. The outer-most loop sets the temperature and the inner-most loop runs a Metropolis Monte Carlo simulation at that temperature. The way in which the temperature is decreased is known as the cooling schedule. In practice, two different cooling schedules are predominantly used; a linear cooling schedule ($T_{new} = T_{old} - dT$) and a proportional cooling schedule ($T_{new} = C \times T_{old}$)

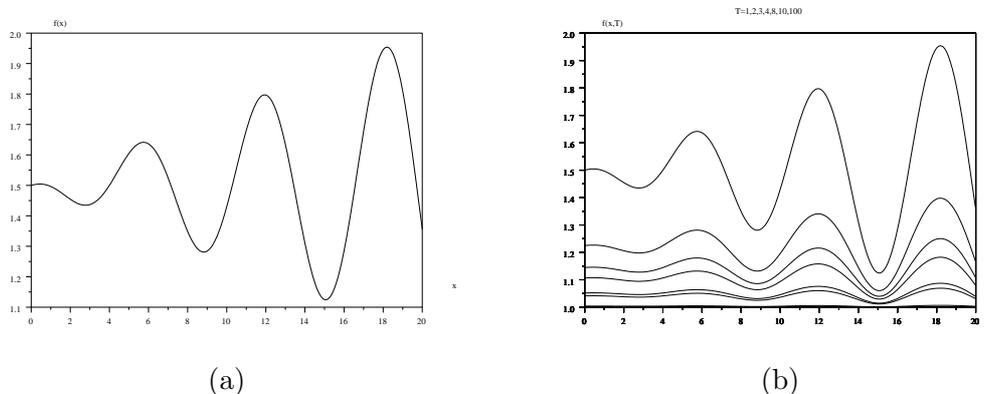


Figure B.1: Simulated annealing principle: the function $f^{1/T}$ (in this case $f(x) = x \cdot \sin(x)$) is sampled for (logarithmically) decreasing temperature in order to locate the global maximum of the function in the early stages of the optimization, concentrating on it in later ones.

where $C < 1.0$. In our implementation instead, the acceptance "temperature" factor is decreased from starting temperature T_0 with iteration step i according to a logarithmic cooling schedule in order to allow for a sufficiently slow confinement of the optimization algorithm to steepest ascent paths.

According to Aarts and Korst (1989), it can be proven that the algorithm converges towards the global minimum if T is lowered not faster than $T^{(n)} = c / \log(n + 1)$, where n denotes the iteration and c is a constant, which usually can be set to one.

Simulated annealing is often considered too computationally expensive to be used in image processing. Since the scene understanding system presented in chapter 6 considers only a limited number of objects in the scene compared to the huge number of image lattice sites, the dimensionality of the problem is reduced, speeding up convergence.

<ul style="list-style-type: none"> • generate initial trial solution trial • best = curr = trial <ul style="list-style-type: none"> – generate trial from curr – if $\text{cost}(\text{trial}) < \text{cost}(\text{best})$: best = curr = trial – else if $\text{cost}(\text{trial}) < \text{cost}(\text{curr})$: curr = trial – else: <ul style="list-style-type: none"> * $\text{anneal} = \exp((\text{cost}(\text{curr}) - \text{cost}(\text{trial})) / \text{temp}(i))$ * generate uniform random number in $[0, 1]$ * if $r < \text{anneal}$: curr = trial
--

Table B.1: Simulated annealing algorithm description.

Bibliography

- E. H. L. Aarts and J. Korst. *Simulated annealing and Boltzmann machines: a stochastic approach to combinatorial optimization and neural computing*. John Wiley and Sons, Tiptree, Essex, 1989.
- F. Adragna, M. Datcu, I. Hajnsek, J. Inglada, K. Papathanassiou, D. Petit, M. Quartulli, and J. C. Souyris. Information extraction from high resolution SAR data. *Proc. XI European Signal Processing Conference, 3-6 sept 2002, Toulouse France EUSIPCO 2002, Paper Num. 855*, Signal and Image Processing for Space Applications (1/2), 2002.
- H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automation and Control*, 19(6):716–723, 1974.
- C. Andrieu, P. Djurić, and A. Doucet. Model selection by MCMC computation. *Signal Processing*, 81:19–37, 2001.
- C. Andrieu, A. Doucet, W. Fitzgerald, and J. Per'ez. Bayesian computational approaches to model selection. In *W.J. Fitzgerald, R.L. Smith, P.C. Young and A. Walden (Eds.) Nonlinear and Non Gaussian Signal Processing*. Cambridge Newton Institute Series, Cambridge University Press, 2000.
- R. Bamler. A comparison of Range–Doppler and Wavenumber domain SAR focussing algorithms. *IEEE Trans. Geosci. and R.S.*, 34(5), 1992.
- R. Bamler and B. Schättler. SAR data acquisition and image formation. In G.Schreier, editor, *SAR Geocoding: Data and Systems*, pages 103–134. Wichmann, Karlsruhe, 1993.
- J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J.Royal Statistical Society B*, 36:192–236, 1974.
- J. Besag. On the statistical analysis of dirty pictures. *J.Royal Statistical Society B*, 48: 259–302, 1986.
- J. Besag, P. Green, D. Higdon, and K. Mengersen. Bayesian computation and stochastic systems (with discussion). *Statistical Science*, 10:3–66, 1995.
- R. E. Blahut. *Principles and practice of information theory*. Addison-Wesley, 1987.
- R. Bolter. Reconstruction of man-made objects from high resolution SAR images. *Proceedings of IEEE Aerospace Conference 2000, Big Sky, Montana, Paper No. 6.0305*, 2000.

- R. Bolter and F. Leberl. Phenomenology-based and interferometry-guided building reconstruction from multiple SAR images. *Proceedings of EUSAR 2000, Munich*, pages 687–690, 2000.
- G. L. Bretthorst. An introduction to model selection using probability theory as logic. In *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder ed. Kluwer Academic Publishers, Dordrecht, 1996.
- P. Brodatz. *Textures: a photographic album for artists and designers*. Dover, New York, 1966.
- S. P. Brooks. MCMC convergence diagnosis via multivariate bounds on log-concave densities. *Annals of Statistics*, 26(1):398–433, 1998.
- G. R. Burkhart, Z. Begen, and R. Carande. Elevation correction and building extraction from interferometric SAR imagery. *IGARSS 96*, pages 659–660, 1996.
- C. Cafforio, C. Prati, and F. Rocca. SAR data focussing using seismic migration techniques. *IEEE Trans. Geosci. and R.S.*, 27(199-207):2, 1991.
- D. Carrasco, D. Esteban, F. López, J. Tena, O. Rodríguez, M. Bara, and A. Broquetas. Precise geometry techniques for ERS SAR interferometry. *Proc. EUSAR'89, Friedrichshafen, Germany*, 1998.
- R. Chellappa. Two-dimensional discrete gaussian Markov random field models for image processing. L. Kanal and A. Rosenfeld editors, *Progress in Pattern Recognition 2, Machine Intelligence and Pattern Recognition*, Elsevier Science Publishers, pages 79–113, 1985.
- R. Chellappa, S. Chatterjee, and R. Bagdazian. Texture synthesis and compression using Gaussian-Markov random field models. *IEEE Trans. Systems Man Cybernet.*, 15(2): 298–303, 1985.
- R. Chellappa and A. Jain. *Markov Random Fields - Theory and Applications*. Academic Press, 1993.
- R. Chellappa, Q. Zheng, L. Davis, C. Liu, X. Zhang, C. Rodriguez, and A. Rosenfeld. Site model based image registration and change detection. Technical report, Contract DACA 76-92-C0024, First Annual Report on Radius Project for the period 30 Sept. 92 – 29 Sept. 93, 1993, 1993.
- F. S. Cohen and D. B. Cooper. Simple parallel hierarchical and relaxation algorithms for segmenting noncausal Markovian random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2(9):195–219, 1987.
- M. Costantini. A novel phase unwrapping method based on network programming. *IEEE Trans. Geosci. and R.S.*, 36(3):813–821, 1998.
- R. T. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13, 1946.

- R. T. Cox. *The algebra of probable inference*. Johns Hopkins Univ Pr., re-published February 2002.
- N. A. C. Cressie. *Statistics for Spatial Data*. Wiley-Interscience, 1991.
- G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(5):25–39, 1983.
- I. Cumming, F. Wong, and R.K.Raney. A SAR processing algorithm with no interpolation. *Proc. of IGARSS'92, Huston, USA*, pages 376–379, 1992.
- I. G. Cumming and J. R. Bennett. Digital processing of SEASAT SAR data. In *IEEE 1979 International Conference on Acoustics, Speech and Signal Processing, (Washington, D.C.), April 2-4, 1979*, 1979.
- J. C. Curlander and R. N. McDonough. *Synthetic Aperture Radar: Systems and Signal Processing*. Wiley, 1992.
- D. A. Yocky D. L. Bickel, W. H. Hensley. The effect of scattering from buildings on interferometric SAR measurements. *IGARSS 97*, pages 1545–1547, 1997.
- M. Datcu, A. Pelizzari, H. Daschiel, K. Seidel, and M. Quartulli. Advanced value adding to metric resolution SAR data: Information mining. *Proc. 4th European Conference on Synthetic Aperture Radar EuSAR2002, 4-6 June 2002, Cologne, Germany*, 2002a.
- M. Datcu and M. Quartulli. Two-level Bayesian inference for modelling and estimation of noisy, nonstationary 2D signals. Submitted for publication on *IEEE Trans. on Image Processing*, 2003.
- M. Datcu and G. Schwarz. The application of forward modelling as a data analysis tool. *Information Processing for Remote Sensing, Ed. C.H. Chen, World Scientific Publishing Company Pte Ltd*, 1999.
- M. Datcu, K. Seidel, S. D Elia, and P.G. Marchetti. Knowledge driven information mining in remote sensing image archives. *ESA Bulletin*, pages 26–33, May 2002b.
- M. Datcu, K. Seidel, and M. Walessa. Spatial information retrieval from remote sensing images - Part I: Information theoretical perspective. *IEEE Trans. Geosci. and R.S.*, 36(5):1431–1445, 1998.
- F. dell'Acqua and P. Gamba. Preparing an urban test site for SRTM data validation. *IEEE Trans. on Geoscience and Remote Sensing*, 40(10):2248–2256, October 2002.
- A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.
- P. J. Diggle. *Statistical Analysis of Spatial Point Patterns*. Academic Press, New York, 1983.
- R. Dubes and A. Jain. Clustering methodologies in exploratory data analysis. *Yovits, M., editor, Advances in Computers, Academic Press*, 19:113–228, 1980.

- Shimon Edelman and Daphna Weinshall. Computational vision: A critical review. Technical Report AIM-1158, MIT Artificial Intelligence Laboratory, 1989.
- C. Elachi. *Spaceborne Radar Remote Sensing: Applications and Techniques*. IEEE, 1988.
- G. A. Fenton. *Simulation and analysis of random fields*. PhD thesis, Princeton University, Princeton, New Jersey, 1990.
- P. Ferrazzoli. SAR for agriculture: Advances, problems and prospects. *Proceedings of LandSAR 2001, 11-14 Sept. 2001, Sheffield, UK*, pages 47–56, 2001.
- W. Fitzgerald, S. Godsill, A. Kokaram, and J. Stark. Bayesian methods in signal and image processing. In *Bayesian Statistics 6*, eds. J.M. Bernardo, et. al., pages 239–254. Oxford University Press, 1999.
- M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker. Query by image and video content: the QBIC system. *IEEE Computer*, 28(9):23–32, 1995.
- M. Fowler. *UML Distilled: A Brief Guide to the Standard Object Modeling Language*. Addison-Wesley Publishing Company, New York, NY, 1999.
- G. Franceschetti, A. Iodice, and D. Riccio. A canonical problem in electromagnetic backscattering from buildings. *IEEE Trans. Geosci. and R.S.*, 40(8):1787–1801, 2002.
- G. Franceschetti and R. Lanari. *Synthetic Aperture Radar Processing*. CRC Press, 1999.
- A. C. Frery, H. J. Müller, C. Yanasse, and S. J. Siqueira. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):648–659, 1997.
- V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman. A model for radar images and its application to adaptive filtering of multiplicative noise. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 4:157–166, 1982.
- U. M. Fyyad and al. Advances in knowledge discovery and data mining. *AAAI Press/ The MIT Press.*, 1996.
- P. Gamba, F. dell’Acqua, and B. Houshmand. SRTM data characterization in urban areas. *Proc. Photogrammetric Computer Vision ISPRS Commission III, Symposium 2002, September 9 - 13, 2002, Graz, Austria*, 2002.
- P. Gamba and B. Houshmand. Three dimensional urban characterization by IFSAR measurements. *IGARSS 99*, pages 2401–2403, 1999.
- P. Gamba and B. Houshmand. Three-dimensional road network by fusion of polarimetric and interferometric SAR data. *Proc. of the 1999 IEEE International Geoscience and Remote Sensing Symposium (IGARSS’99)*, 1:302–304, Hamburg, Germany, June 1999.

- E. Gamma, R. Helm, R. Johnson, and J. Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley Professional Computing Series. Addison-Wesley Publishing Company, New York, NY, 1995.
- F. Gatelli, A. Monti Guarnieri, F. Parizzi, P. Pasquali, C. Prati, and F. Rocca. The wavenumber shift in SAR interferometry. *IEEE Trans. Geosci. and R.S.*, 32(4), 1994.
- D. Geman. Random fields and inverse problems in imaging. *Lecture Notes in Mathematics*, Springer, pages 117–193, 1998.
- S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.
- C. Geyer. Likelihood inference for spatial point processes. *Skandinavian journal of statistics*, 21:359–373, 1994.
- D. C. Ghiglia and L. A. Romero. Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America*, 11(1), 1998.
- R. M. Goldstein and C. M. Werner. Radar ice motion interferometry. *Proc. of the 3rd ERS Symposium, Florence, Italy*, 1997.
- J. W. Goodman. Statistical properties of laser speckle patterns. In J.C.Dainty, editor, *Laser speckle and related phenomena*. Springer-Verlag, Berlin, 1975.
- L. Van Gool, P. Dewaele, and A. Oosterlicck. Texture analysis. *CVGIP*, 29:336–357, 1985.
- S. F. Gull. Bayesian inductive inference and maximum entropy. In G.J.Erickson and C.R.Smith, editors, *Maximum Entropy and Bayesian methods in science and engineering*, volume 1: Foundations. Kluwier, 1988.
- D. Heckerman, D. Geiger, and D. M. Chickering. Learning Bayesian networks: the combination of knowledge and statistical data. *Machine Learning*, 20(3):197–243, 1995.
- E. Haines. Point in polygon strategies. *in: Graphics Gems IV, edited by P.S.Heckelbert, Academic Press*, pages 24–46, 1994.
- R. M. Haralick, K. Shanmugan, and I. Dinstein. Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics*, 3(6):610–621, Nov. 1973.
- D. Heckerman. A tutorial on learning with Bayesian networks. Technical Report MSR-TR-95-06, Microsoft Research, Redmond, Washington, 1995 Revised June 96., 1995.
- D. Heckerman, D. Geiger, and D. M. Chickering. Learning Bayesian networks: The combination of knowledge and statistical data. In *KDD Workshop*, pages 85–96, 1994.
- S. Heuel and R. Nevatia. Including interaction in an automated modelling system. *IEEE Symposium on Computer Vision*, pages 383–388, 1995.

- R. Huber and L. V. Dutra. Feature selection for ERS1/2 InSAR classification: High dimensionality case. *IGARSS 98*, 1998.
- A. Huertas, Z. Kim, and R. Nevatia. Use of IFSAR with intensity images for automatic building modeling. *DARPA Project, Annual Report 1998, Los Angeles, California*, 1998.
- M. Imberty and X. Descombes. Simulation de processus objets: Etude de faisabilité pour une application à la segmentation d'image. *INRIA, Rapport De Recherche N.3881, France, February 2000*, 2002.
- E. T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106(620), 1957.
- E. T. Jaynes. Prior probabilities. *IEEE Trans. on Systems Science and Cybernetics*, sec-4 (3):227–241, 1968.
- E. T. Jaynes. On the rationale of maximum entropy methods. *Proc. IEEE*, 70(9), 1982.
- E. T. Jaynes. Bayesian methods: general background. In J.H. Justice, editor, *Maximum Entropy and Bayesian methods in applied statistics*. CUP, 1986.
- E. T. Jaynes. Bayesian spectrum and chirp analysis. *Maximum Entropy and Bayesian Spectral Analysis and Estimation Problems*, C. R. Smith and G. J. Erickson eds., page 1, 1987.
- E. T. Jaynes. *Probability Theory - The Logic of Science - G. L. Bretthorst (Ed.)*. Cambridge University Press, 2003.
- H. Jeffreys. *Theory of Probability*. Oxford Univ. Press, 1939.
- M. Y. Jin and C. Wu. A SAR correlation algorithm which accomodates large range migrations. *IEEE Trans. Geosci. and R.S.*, 22:592–597, 1984.
- S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220:671–680, 1983.
- R. Kothari and D. Ensley. On the use of model order for detecting potential target locations in SAR images. *Proc. SPIE on Algorithms for Synthetic Aperture Radar Imagery V*, 3370: 351–357, 1998.
- R. Lanari. A new method for the compensation of the SAR range cell migration based on the Chirp-z transform. *Trans. Geosci. and Remote Sensing*, 33(5):1229–1299, 1995.
- R. Lanari and G. Fornaro. A short discussion on the exact compensation of the SAR range-dependent range cell migration effect. *Trans. Geosci. and Remote Sensing*, 35(6): 1446–1452, 1997.
- R. Lanari, S. Hensley, and P. A. Rosen. Chirp-z transform based SPECAN approach for phase-preserving ScanSAR image generation. *Proc. Instr. Electr. Eng. Radar Sonar Navig.*, 145:254–261, 1998.

- H. Laur, S. Coulson, M. Doherty, and N. Walker. ERS SAR processing and post-processing at ESA. In *Proc. of CEOS SAR Calibration Workshop*, 1998.
- J Lee, K. P. Papathanassiou, T. L. Ainsworth, M. R. Grunes, and A. Reigber. A new technique for noise filtering of SAR interferometric phase images. *IEEE Trans. Geosci. and R.S.*, 36(5), 1998.
- J. S. Lee. Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, 17:24–32, 1981.
- J. S. Lee. A simple speckle smoothing algorithm for synthetic aperture radar images. *IEEE Transactions ON Systems, Man and Cybernetics*, SMC-13(17):85–89, 1983.
- J. S. Lee. Speckle suppression and analysis for synthetic aperture radar images. *Optical Engeneering*, 25(5):636–643, 1986.
- J. S. Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller. Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Trans. Geosci. and R.S.*, 32(5):1017–1028, 1994.
- J. Li, R. Nevatia, and S. Nornoha. User assisted modeling of buildings from aerial images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, II:274–279, 1999.
- S. Z. Li. *Markov random field modeling in Computer Vision*. T. L. Kunii (Ed.), Springer-Verlag, Tokyo, 1995.
- C. Lin and R. Nevatia. Building detection and description from a single intensity image. *Computer Vision and Image Understanding*, pages 101–121, 1998.
- O. Loffeld, A. Hein, and F. Schneider. SAR focusing: scaled inverse Fourier transformation and chirp scaling. In *Proc. Int. Geosci. and Remote Sensing Symp., June 1998*, pages 630–632, 1998.
- A. Lopes, E. Nezry, R. Touzi, and H. Laur. Structure detection and statistical adaptive speckle filtering in SAR images. *International Journal of Remote Sensing*, 9(14):1735–1758, 1993.
- D. J. C. MacKay. Maximum entropy and Bayesian methods in inverse problems. in *Bayesian Interpolation*, Kluwer Academic Publisher, pages 39–66, 1991.
- D. J. C. MacKay. Bayesian interpolation. *Neural Computation*, 4(3):415–447, 1992.
- D. J. C. MacKay. *Information theory, inference and learning algorithms*. Cambridge University Press, 2002.
- D. Marr. *Vision*. W. H. Freeman, San Francisco, CA, 1982.
- J. L. Marroquin and F. Girosi. Some extensions of the K-means algorithm for image segmentation and pattern classification. *AIM*, 1390:21, 1993.

- D. Massonnet. Geoscientific applications at cnes. In G.Schreier, editor, *SAR Geocoding: Data and Systems*, pages 103–134. Wichmann, Karlsruhe, 1993.
- H. Mayer. Automatic object extraction from aerial imagery - a survey focusing on buildings. *Computer Vision and Image Understanding*, 74(2):138–149, 1999.
- I. McConnell and C. Oliver. Comparison of annealing and iterated filters for speckle reduction in SAR. *Microwave Sensing and Synthetic Aperture Radar*, edited by G. Franceschetti et al, *Proc. SPIE*, 2958:74–85, 1996.
- M. E. Mejail, A. C. Frery, J. Jacobo-Berlles, and O. H. Bustos. Approximation of distributions for SAR images: proposal, evaluation and practical consequences. *Latin American Applied Research*, 31:83–92, 2001.
- N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J.of Chem. Phys.*, 21(6):1087–1092, 1953.
- Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey. <http://terraserver-usa.com/>, Baltimore map 01 jul 1982.
- Microsoft Terraserver Imagery, Images courtesy of the U.S. Geological Survey. <http://terraserver-usa.com/>, LasVegas aerial photographs, 10 jun 1994.
- T. P. Minka and R. W. Picard. Interactive learning with a 'society of models'. *Pattern Recognition*, 30(4):565–581, 1997.
- T. K. Moon. The Expectation-Maximization algorithm. *IEEE Signal Processing Magazine*, 11:47–60, 1996.
- P. Myllymaki and H. Tirri. Bayesian case-based reasoning with neural networks. In *Proc. of the IEEE International Conf. on Neural Networks, San Francisco, March 1993*, pages 422–427. IEEE Press, 1993.
- G. Nico, M. Walessa, and M. Datcu. Topographic mapping: a new approach. *Proc. IGARSS'98, Seattle, Washington, Jul. 1998*, 1:91–93, 1998.
- A. O'Hagan. *Kendall's Advanced Theory of Statistics*, volume 2B Bayesian Inference. Arnold, London, 1994.
- C. J. Oliver. Information from SAR images. *J. Phys. D: Appl. Phys.*, 24:1493–1514, 1991.
- C. J. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar images*. Artech House, Boston and London, 1998.
- J. O'Rourke. *Computational Geometry in C*. Cambridge University Press, 1998.
- M. Ortner, X. Descombes, and J. Zerubia. Building extraction from digital elevation model. *INRIA, Rapport De Recherche N.4517, France, July 2002*, 2002.
- A. Papoulis. *Signal analysis*. MacGraw-Hill, New York, 1977.

- A. Papoulis. *Probability, random variables and stochastic processes*. MacGraw-Hill, Singapore, 1984.
- V. Pavlovic, R. Sharma, and T. S. Huang. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):677–695, 1997.
- A. Pen, J. Lozano, and J. Larran. An empirical comparison of four initialization methods for the k-means algorithm. *Pattern Recognition Letters*, 20(50):1027–1040, 1999.
- A. Piardi, H. Daschiel, M. Quartulli, S. Serpico, F. Tupin, and M. Datcu. Scene understanding for settlements from metric resolution imagery. *Proc. IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas URBAN2001, Rome, Italy, 8-9 nov 2001*, 2001.
- M. Pincus. A Monte Carlo method for the approximate solution of certain types of constrained optimization problems. *Operative Research*, 18:1225–1228, 1970.
- C. Prati, F. Rocca, A. Monti Guarnieri, and P. Pasquali. Report on ERS-1 SAR interferometric techniques and applications. Technical report, ESA Study Contract Report , ESA Contract N.:3-7439/92/HGE-I, 1994.
- RADARSAT-2 presentation web site. <http://www.radarsat2.info/>, 1995-2003.
- M. Quartulli and M. Datcu. A Bayesian inversion method for the 3D reconstruction of settlements from metric SAR observations. *Proceedings of SPIE 2001, 17-18 sept. 2001, Toulouse, France*, 4543 SAR Image Analysis, Modeling and Techniques IV:161–168, 2001a.
- M. Quartulli and M. Datcu. Bayesian model based city reconstruction from high resolution InSAR data. *Proc. IEEE/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas URBAN2001, Rome, Italy, 8-9 nov 2001*, 2001b.
- M. Quartulli and M. Datcu. Quality of model-based information extraction from SAR images. *Proceedings of SPIE 2001, 17-18 sept. 2001, Toulouse, France*, 4543 SAR Image Analysis, Modeling and Techniques IV, 2001c.
- M. Quartulli and M. Datcu. Scene understanding for settlements from metric resolution imagery. *Proc. 3rd Int. Symposium on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications LandSAR2001, 11-14 Sept.2001, Sheffield, UK*, 2001d.
- M. Quartulli and M. Datcu. 3-d scene reconstruction and structure recognition from metric resolution InSAR. *Proc. 4th European Conference on Synthetic Aperture Radar EuSAR2002, 4-6 June 2002, Cologne, Germany*, 2002.
- M. Quartulli and M. Datcu. Information extraction from high resolution SAR data for urban scene understanding. In *Proc. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas Urban2003, Berlin, Germany, 22-23 may 2003*, 2003a.

- M. Quartulli and M. Datcu. Information fusion for scene understanding from interferometric SAR data in urban environments. *IEEE Transactions on Geoscience and Remote Sensing*, 41(9):1976–1985, 2003b.
- M. Quartulli and M. Datcu. Structural information extraction from high resolution SAR data. In *Proceedings of Interational Geoscience And Remote Sensing Symposium IGARSS 2003, Toulouse France 20-26 july 2003*, 2003c.
- M. Quartulli and M. Datcu. Structural object modeling for high resolution SAR urban scene understanding. In *Proc. 2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas Urban2003, Berlin, Germa, 22-23 may 2003*, 2003d.
- M. Quartulli and M. Datcu. Stochastic geometrical modeling for urban scene understanding from a single SAR intensity image with meter resolution. *IEEE Trans. Geosci. and Remote Sensing*, 42(9), 2004.
- R. K. Rainey and G. J. Wessels. Spatial considerations in SAR speckle simulation. *IEEE Transactions on Geoscience and Remote Sensing*, 26(5):666–671, 1988.
- T. Randen and J. Husoy. Filtering for texture classification: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 4(21):291–309, 1999.
- R. K. Raney, H. Runge, R. Bamler, I. G. Cumming, and F. H. Wong. Precision SAR processing using chirp scaling. *IEEE Trans. Geosci. and R.S.*, 32(4), 1994.
- A. W. Rihaczek and S. J. Hershkowitz. *Radar resolution and complex-image analysis*. Artech House, 1996.
- B. D. Ripley. Modelling spatial patterns. *Journal of the Royal Statistical Society B*, 39: 172–192, 1977.
- B. D. Ripley and F. P. Kelly. Markov point processes. *J. London Math. Soc.*, 15:188–192, 1977.
- J. Rissanen. Modeling by shortest data description. *Automatica*, 1978.
- J. Rissanen. Minimum-description-length principle. In *Encyclopedia of Statistical Sciences*, volume 5, pages 523–527. Wiley, New York, 1985.
- F. Rocca, C. Prati, and A. Monti Guarnieri. New algorithms for processing of SAR data. Technical report, ESA-ESRIN contract, 1989.
- D. Ruelle. *Statistical mechanics: rigorous results*. Benjamin, New York, 1969.
- H. Runge and R. Bamler. A novel high precision SAR focussing algorithm based on chirp scaling. *Proc. of IGARSS'92, Huston, USA*, pages 372–375, 1992.
- K. Sarabandi. Derivation of phase statistics from the Mueller matrix. *Radio Science*, 27(5): 553–560, 1992.

- R. Scheiber, A. Reigber, A. Ulbricht, K. P. Papathanassiou, R. Horn, S. Buckreuß, and A. Moreira. Overview of interferometric data acquisition and processing modes of the experimental airborne SAR system of DLR. *Proc. IGARSS'99, Hamburg, Germany*, 1999.
- G. Schreier. Geometrical properties of SAR images. In G. Schreier, editor, *SAR Geocoding: Data and Systems*, pages 103–134. Wichmann, Karlsruhe, 1993.
- M. Schröder. Interactive learning in remote sensing image databases. *IEEE Intern. Geoscience and Remote Sensing Symposium IGARSS 99*, 1999a.
- M. Schröder. *Stochastic Modeling of Image Content in Remote Sensing Image Archives*. PhD thesis, ETH Zurich Switzerland, 1999b.
- M. Schröder and H. Rehrauer. MMDemo documentation. *Image Science Group, ETH Zurich*, 1999.
- M. Schröder, H. Rehrauer, K. Seidel, and M. Datcu. Spatial information retrieval from remote sensing images - Part II: Gibbs Markov fields. *IEEE Trans. Geosci. and R.S.*, 36(5):1446–1455, 1998a.
- M. Schröder, H. Rehrauer, K. Seidel, and M. Datcu. Interactive learning and probabilistic retrieval in remote sensing image archives. *IEEE Transactions on Geoscience and Remote Sensing*, 38:2288–2298, 2000a.
- M. Schröder, K. Seidel, and M. Datcu. Gibbs random field models for image content characterization. *Stein, T. I., editor, IEEE International Geoscience and Remote Sensing Symposium (IGARSS97)*, 1:258–260, 1998b.
- M. Schröder, M. Walessa, H. Rehrauer, K. Seidel, and M. Datcu. Gibbs random field models: A toolbox for spatial information extraction. *Computers and Geosciences*, 26(4):423–432, 2000b.
- M. Schwäbisch. DFD-DLR X-SAR unwrapped Mount Etna, 8 and 9 October 1994 Shuttle data, <http://ceos.cnes.fr:8100/cdrom-00b2/ceos1/casestud/fdb/aetna2.htm>, 1994.
- K. Seidel, M. Schröder, H. Rehrauer, and M. Datcu. Query by image content from remote sensing archives. *IEEE Intern. Geoscience and Remote Sensing Symposium IGARSS98*, 1:393–396, 1998.
- C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.
- E. Simonetto. *Extraction 3-D de Structures Industrielles sur des Images RAMSES Haute Resolution par Radargrammetrie*. PhD Thesis, Universite de Rennes 1, 2002.
- E. Simonetto and H. Oriot. Analyse du contenu topographique d'images radar haute resolution dans un contexte radargrammtrique. *Presentation au GDR-ISIS, reunion PR2, ENST Paris*, 2000.
- E. Simonetto, H. Oriot, and R. Garello. 3d extraction from airborne SAR imagery. *Proc. EUROPTO'99, Florence*, 1999.

- E. Simonetto, H. Oriot, and R. Garello. Extraction of industrial structures and DEM from airborne SAR images. *Proceedings of the Second International Symposium, PSIP'2001, Marseille, 23-24 January 2001, France*, pages 412–417, 2001.
- D. S. Sivia. *Data analysis: a Bayesian tutorial*. Clarendon Press, Oxford, 1996.
- A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(12): 1349–1380, 2000.
- S. Stan, G. Palubinskas, and M. Datcu. Bayesian selection of the neighborhood order for Gauss-Markov texture models. *Elsevier Pattern Recognition Letters*, 23:1229–1238, 2001.
- U. Stilla, U. Sörgel, and U. Thönnessen. Geometric constraints for building reconstruction from InSAR data of urban areas. *Proc. Symposium on geospatial theory, processing and applications. International Archives of Photogrammetry and Remote Sensing*, 34(4), 2002.
- U. Stilla, U. Sörgel, and U. Thönnessen. Potential and limits of InSAR data for building reconstruction in built-up areas. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58:113–123, 2003.
- R. Stoica, X. Descombes, and J. Zerubia. Road extraction in remote sensed images using a stochastic geometry framework. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 20th International Workshop Gif-Sur-Yvette, France 8-13 July 2000*, 2000.
- D. Stoyan and H. Stoyan. On one of Mathern's hard-core point process models. *Mathematische Nachrichten*, 122:205–214, 1985.
- R. Touzi. A protocol for speckle filtering of SAR images. *Proc. CEOS'99, Toulouse, France, 26 - 29 October, 1999*.
- R. Touzi and A. Lopes. Statistics of the Stokes parameters and of the complex coherence parameters in one-look and multilook speckle fields. *IEEE Trans. Geosci. and R.S.*, 34(2):519–531, 1996.
- R. Touzi, A. Lopes, J. Bruniquel, and P. W. Vachon. Coherence estimation for SAR imagery. *IEEE Trans. Geosci. and R.S.*, 37(1):135–149, 1999.
- H. L. Van Trees. *Detection, Estimation, and Modulation Theory*. Wiley, 1968.
- M. Tuceryan and A. Jain. Texture segmentation using Voronoi polygons. *IEEE PAMI*, 12: 211–216, 1990.
- F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky. Detection of linear features in SAR images: Application to road network extraction. *IEEE Trans. Geosci. and R.S.*, 36(2):434–453, 1998.
- F. Tupin and M. Roux. Detection of building outlines based on the fusion of SAR and optical features. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58:71–82, 2003.

- M. Tur, K. C. Chin, and J. W. Goodman. When is speckle noise multiplicative? *Applied Optics*, 21(7):1157–1159, 1982.
- G. J. G. Upton and B. Fingleton. *Spatial Data Analysis by Example, Volume 1, in Point Pattern and Quantitative Data*. Wiley, Chirchester, 1985.
- S. Vanhuyse, E. Wolff, M. Acheroy, Y. Yvinec, D. Borghys, H. Suss, M. Keller, M. Bajuic, I. Bloch, Y. Yu, and O. Damanet. Space and airborne mined area reduction tools - poster. In *Space and Disasters Management forum, 26 March 2003*, 2003.
- M. Walessa. *Bayesian Information Extraction from SAR Images*. PhD Thesis, Universitaet Siegen, 2000.
- M. Walessa and M. Datcu. Model-based despeckling and information extraction from SAR images. *IEEE Trans. Geosci. and R.S.*, 38(5):2258–2269, 2000.
- SRTM web site. <http://www.dfd.dlr.de/srtm/index.html>. DLR DFD web site, 2000.
- G. Winkler. *Image Analysis, Random Fields and Dynamic MonteCarlo Methods*. Springer-Verlag, 1995.
- I. H. Woodhouse. Stop, look and listen: auditory perception analogies for radar remote sensing. *International Journal of Remote Sensing*, 21(15):2901–2913, 2000.
- C. Wu, K. Y. Liu, and M. Jin. Modelling and a correlation algorithm for spaceborne SAR signals. *IEEE Trans. Geosci. and R.S.*, 18(5):563–574, 1982.
- J. S. Yedidia, W. T. Freeman, and Y. Weiss. Understanding belief propagation and its generalizations. in *Exploring Artificial Intelligence in the New Millennium*, eds. G. Lakemeyer and B. Nebel, Morgan Kaufmann, pages 239–269, 2003.

Curriculum vitae

Professional Carrer:

november 2000 till date:

DLR IMF, Oberpfaffenhofen Germany
Working on model-based high resolution SAR and InSAR information extraction and data mining systems with the Image Analysis group of Prof.Dr.Mihai Datcu at the German Aerospace Center DLR in the Institute for Remote Sensing Technology, IMF directed by Prof.Dr.Richard Bamler

october 1998 till date:

ACS, Rome Italy
Working in Rome on SAR and InSAR data processing systems research and development in the Earth Observation division of Advanced Computer Systems, an Italian company producing satellite data acquisition and processing chains at receiving ground stations, digital image analysis systems, remote sensing applications and multimedia digital archives

Education:

till october 1998:

University of Bari, Italy
'Laurea' degree in Physics with 107/110 marks on the 20th of october with a thesis on "SAR Differential Interferometry: Theory and Applications" ("Teoria e applicazioni dell'interferometria SAR differenziale") under the supervision of Prof.Dr. Luciano Guerriero

till 1992:

Classical studies Lyceum in Bari, 60/60 exit mark