

Multichannel SAR Interferometry based on Statistical Signal Processing



Giampaolo Ferraioli

Università degli Studi di Napoli Parthenope

Dipartimento per le Tecnologie

A thesis submitted for the degree of

Philosophiæ Doctor (PhD) in Telecommunication Engineering

December, 2008

Università degli Studi di Napoli “Parthenope”
Dipartimento per le Tecnologie

**Multichannel SAR Interferometry based on Statistical Signal
processing**

Giampaolo Ferraioli

A thesis submitted for the degree of
Philosophiæ Doctor (PhD)
in
Telecommunication Engineering

Promoter

Prof. Vito Pascazio

Università degli Studi di Napoli Parthenope, Napoli, Italy

Jury

Prof. Adriano Jose Camps Carmona (Chairman)

Universitat Politecnica de Catalunya, Barcelona, Spain

Prof. Maurizio Migliaccio

Università degli Studi di Napoli Parthenope, Napoli, Italy

Prof. Lorenzo Bruzzone

Università degli Studi di Trento, Trento, Italy

Napoli, 15 December 2008

Abstract

Interferometric SAR systems allow to reconstruct height profile of earth surfaces. The height reconstruction is based on phase unwrapping operation, which is an ill-posed problem since it admits infinite number of solutions. The phase unwrapping problem can be solved using the multichannel InSAR approach, based on the combination of different images of the scene obtained by slightly different positions. An effective way to combine the different interferograms is using statistical estimation theory. In particular the Maximum Likelihood and Maximum a Posteriori multichannel approach have proved to be effective and to be able to restore the uniqueness of the solution. In this thesis the statistical multichannel phase unwrapping is deeply analysed. In particular, instruments and methods to use InSAR multichannel configuration on real data and in urban areas are provided. Moreover, a new fast and efficient multichannel phase unwrapping algorithm is provided. The presented innovative techniques have been tested on simulated data and real data, showing the consistence and effectiveness of the proposed approaches.

Sommario

I sistemi Interferometrici SAR permettono la ricostruzione di profili di quota della superficie terrestre. La ricostruzione di quota è basata sullo srotolamento della fase. Il problema dello srotolamento della fase è un problema mal posto in quanto ammette un numero infinito di soluzioni. Tale problema può essere risolto utilizzando l'approccio interferometrico multicanale, basato sulla combinazione di diverse immagini della stessa scena, ottenute da posizioni leggermente diverse. Un metodo efficace per combinare i diversi interferogrammi è fornito dalla teoria della stima statistica. In particolare, è stato dimostrato che gli stimatori Maximum Likelihood e Maximum a Posteriori sono efficaci e sono in grado di restaurare l'unicità della soluzione. In questo lavoro di tesi lo srotolamento di fase statistico multicanale è esaminato a fondo. In particolare, sono forniti strumenti e metodi per utilizzare la configurazione interferometrica multicanale su dati reali e in ambiente urbano. Infine, un nuovo algoritmo veloce ed efficiente per lo srotolamento multicanale della fase è mostrato. Le innovative tecniche presentate sono state testate su dati simulati e su dati reali, mostrando la consistenza e l'efficacia degli approcci proposti.

*to Rosa and Luisa
and to Whom, in spite of all,
has always relied on me*

Acknowledgements

I would like to acknowledge all the people that directly or indirectly have contributed to this thesis.

I wish to thank...

...Prof. Vito Pascazio for being, once again, my thesis tutor. Don't worry Vito, this should be my last thesis...but never say never.

...Prof. Florence Tupin, for giving me the opportunity to spend six months working with her and her research group. It has been a great experience, both under a professional and a human point of view.

...Prof. Maurizio Migliaccio, for being an excellent PhD School Director.

...Fabio Baselice, Aymen Shabou and Giancarlo Ferraiuolo: colleagues and friends at the same time.

...Dr. Gianfranco Fornaro, from IREA, for his interesting and useful hints.

...Dr. Andreas Reigber, from DLR, for providing the E-SAR data and for his useful suggestions.

...Paolo Cuccuru, Antonio Sorrentino, Roberta Autieri, Roberto Ambrosino, Attilio Gambardella, Sara Cacciapuoti and Marcello Caleffi that in different ways took part in this experience.

...Ferdinando Nunziata...we have been standing each other for eight long years.

...my parents, Francesco, Enzo and Andrea, for being comprehensive and patient.

Last but not least, I wish to thank Teresa for sharing this “adventure” with me.

Contents

List of Figures	ix
List of Tables	xi
1 Introduction	1
2 Statistical Signal Processing	5
2.1 Introduction	5
2.2 Labeling Problems	5
2.3 Classical Statistical Estimation Theory	7
2.4 Bayesian Estimation Theory	10
2.4.1 Markov Random Fields	13
2.4.1.1 Sampling an MRF	15
2.5 MAP-MRF framework	18
2.5.1 Parameter Estimation	19
2.5.1.1 Complete data problem	19
2.5.1.2 Incomplete data problem	21
2.5.2 Optimization in the MAP-MRF framework	22
2.5.2.1 Simulated Annealing	22
2.5.2.2 Iterated Conditional Modes	24
2.6 Conclusions	24
3 Multichannel SAR Interferometry	27
3.1 Introduction	27
3.2 Synthetic Aperture Radar	28
3.2.1 SAR images	30

CONTENTS

3.2.2	SAR processing	32
3.3	SAR Interferometry	33
3.3.1	Interferometer Height Sensibility	35
3.3.2	Statistical Description	37
3.3.2.1	Coherence Coefficient	38
3.3.3	Phase Unwrapping	40
3.4	Multichannel SAR Interferometry	42
3.4.1	Maximum Likelihood Estimation	45
3.4.2	Maximum a Posteriori Estimation	48
3.5	Conclusions	51
4	Phase Offset Estimation in Multichannel SAR Interferometry	55
4.1	Introduction	55
4.2	Phase Offset in Multichannel InSAR Systems	56
4.3	Phase Offset Estimation	57
4.4	Simulated Experiments	60
4.5	Real Data Experiments	63
4.6	Conclusions	65
5	Multichannel InSAR in Urban Areas	67
5.1	Introduction	67
5.2	InSAR Layover Phenomenology in Urban Scenario	68
5.3	Urban Edge Detection	70
5.3.1	InSAR Urban Edge Detection	73
5.3.1.1	Numerical Experiments on Simulated and Real Data	74
5.4	MAP multichannel DEM reconstruction using Auxiliary Data	79
5.4.1	Simulated Experiments	82
5.5	Conclusions	85
6	Fast InSAR Multichannel Phase Unwrapping	87
6.1	Introduction	87
6.2	Fast MAP Multichannel Phase Unwrapping	88
6.3	Local Energy function	89
6.4	Graph-cut Based Optimization	91

CONTENTS

6.4.1	Ishikawa algorithm	92
6.4.2	Alpha-expansion algorithm	95
6.4.3	Hyperparameter estimation	97
6.5	Simulated Experiments	98
6.6	Real Data Experiments	101
6.7	Conclusions	102
7	Further Works	105
7.1	Introduction	105
7.2	Multichannel SAR Tomography	105
7.3	Ground Based SAR	108
8	Conclusion	113
	References	115

CONTENTS

List of Figures

2.1	Cost Functions	11
2.2	Neighborhood Systems and Cliques	14
3.1	Radar acquisition geometry	28
3.2	SAR acquisition system	30
3.3	SAR geometrical distortions	31
3.4	InSAR geometry	34
3.5	InSAR height Sensitivity	36
3.6	Interferometric phase pdf	39
3.7	Multifrequency interferograms generation	43
3.8	Multifrequency configuration	44
3.9	Multibaseline configuration	44
3.10	Multichannel Likelihood Function	47
4.1	High Coherence Pixels	58
4.2	Phase offset simulated experiment - ramp profile	61
4.3	Ramp profile results	62
4.4	Phase offset simulated experiment - Gaussian profile	63
4.5	Gaussian profile results	64
4.6	Phase offset real data experiment	66
5.1	Layover Building Geometry	68
5.2	Buildings in Layover Areas	70
5.3	Edges in Layover areas	72
5.4	InSAR Edge detection	75
5.5	Layover Area Edge detection	77

LIST OF FIGURES

5.6	Real Data Edge detection	78
5.7	Optical and SAR geometry	80
5.8	DEM reconstruction in Layover areas - data	82
5.9	DEM reconstruction in Layover areas - results	84
6.1	Graph representation	92
6.2	Ishikawa Graph Construction	93
6.3	Ishikawa edges weights	94
6.4	α -expansion graph	96
6.5	Comparison between ICM-LGMRF and the GC-TV approach	99
6.6	Comparison between α -expansion and Ishikawa algorithm.	101
6.7	3D reconstruction of real data	102
7.1	Layover Geometry	106
7.2	Interferometric GB-SAR system	109
7.3	GBSAR Interferometry: CRLBs behaviour	111
7.4	GBSAR Interferometry: CRLBs vs coherence	112

List of Tables

4.1	Phase Offset Estimation, Ramp-like Profile	61
4.2	Phase Offset Estimation, Gaussian Profile	62
6.1	α -expansion edge weights	97
6.2	Comparison between α -expansion and Ishikawa	100

LIST OF TABLES

1

Introduction

Statistical Signal Processing provides useful instruments to solve many problems. In particular, Statistical Signal Processing is widely used to analyse and solve labeling problems that occur in computer vision.

A labeling problem consists in finding the optimal configuration of image pixels (i.e. attribution of labels to pixels) according to a specific criterion. Due to the uncertainties in vision problem, the labeling problem is defined according to an optimal criterion.

The best way to find the optimal solution for the labeling problem is to exploit the instruments provided by the statistical estimation theory. There exist two branches of the statistic estimation theory. The first one is known as Classical Estimation, the second one is the known as Bayesian Estimation. This two branches give arise to two optimal criteria widely used: respectively the maximum likelihood and maximum a posteriori criterion.

In particular, one of the most popular instruments provided by the statistical estimation is the so-called *MAP-MRF* framework, that arises from the union of Markov Random Fields theory and Bayesian estimation.

In this thesis statistical estimation instruments are applied to *Synthetic Aperture Radar* (SAR) image processing, in the interferometric configuration (InSAR). InSAR technique allows the reconstruction of height profiles of earth surface. The main task of InSAR processing is to solve the *Phase Unwrapping* (PhU) problem. This is an ill posed problem if the so called *Itoh condition* is violated. An efficient approach to solve the PhU problem is to use MAP-MRF framework together with

1. INTRODUCTION

a multifrequency and or a multibaseline SAR raw datasets. In this approach a *Local Gaussian MRF* (LGMRF) is used to model the a priori distribution and the *Iterated Conditional Modes* (ICM) algorithm is used for the optimization step.

The aim of this thesis is to deeply study and analyze the InSAR processing based on Statistical Signal Processing. Recalls of Statistical Estimation theory are made in the second chapter, with a particular attention to Markov Random Field theory. Interferometric SAR systems and the height reconstruction problem are addressed in the third chapter.

In the fourth chapter of the thesis the problem of the phase offset on interferometric data is considered and investigated. Real multichannel interferometric data are known except for not controllable phases. This phase offset results as the combination of various factors such as a nonperfect synchronization in the acquisition system, SAR processing errors, atmospheric effects, and parallel-baselines uncertainty. The phase offset is different for each different interferogram, but it can be reasonably supposed to be slowly varying or spatially homogeneous on a single interferogram. Estimation of these offsets is mandatory in order to retrieve the absolute height profile using the multichannel LGMRF-ICM approach. A technique to estimate phases offset, based on ML estimation is proposed. The effectiveness of this approach on height surface reconstruction problem is validated and assessed on simulated data and real data.

In the fifth chapter the attention is focused on multichannel InSAR systems in urban scenario. First an approach based on multichannel InSAR systems for the edge detection in urban areas is proposed. The algorithm works on the difference height between pixels. It is based on the estimation of the local hyperparameters characterizing the Local Gaussian MRF (LGMRF) used to model the unknown pixels height. The hyperparameters can be seen as an index of spatial correlation between adjacent pixels. The method is tested on simulated and real data, showing its capacity to overcome the problems typical of SAR edge detection, due to the multiplicative behaviour of speckle. Then, an improvement of the MAP multichannel phase unwrapping technique in case of urban layovered areas is proposed. A methodology exploiting auxiliary data in DEM reconstructions is applied, using optical data or the shadowing information from the SAR image.

These additional information are used in order to improve the a priori information, from which the MAP multichannel height estimation starts. The results obtained on simulated data show the goodness of the method.

In the sixth chapter a new fast algorithm, based on the MAP-MRF framework, for the multichannel phase unwrapping problem is proposed. The idea is to improve the LGMRF-ICM multichannel phase unwrapping. Despite its effectiveness especially dealing with noisy data and big discontinuities, the LGMRF-ICM approach can be excessively time consuming and computationally heavy due to both a priori model and optimization step. The proposed idea is to overcome these limits by introducing a new a priori model based on Total Variation and using energy optimization algorithms based on graph-cut theory. The new approach gives similar solutions to the previous one but in much less computation time. Furthermore, the exact solution of the phase unwrapping problem, related to the new model, can be provided. A set of experimental results on both simulated and real data illustrates the effectiveness of the approach.

Finally a quick overview on future developments of this work are shown. First SAR Tomography is investigated in order to separate and to discriminate the three complex contributions which collapse in layover pixels. A new solution based on statistical signal processing and multichannel systems for the previous problem is proposed. Then, Ground Based Interferometric systems are analysed. Preliminary studies on the achievable height reconstruction accuracy using a multichannel configuration in the Ground Based systems are conducted.

1. INTRODUCTION

2

Statistical Signal Processing

2.1 Introduction

In this first chapter the labeling problem in computer vision is addressed. A labeling problem consist into assigning a *label* configuration to a set of *sites*. Due to the uncertainties in vision problem, the labeling solution is defined according to an optimal criteria. The most used optimal criteria used in vision problems are provided by the Maximum Likelihood Estimation and the Maximum a Posteriori Estimation. The latter, together with *Markov Random Field* theory gives arise to the MAP-MRF framework, which is a convenient and consistent instrument to solve the mentioned problem. In the first part of the chapter, after introducing the labeling problem, recalls of Statistical Estimation theory are made, related to both Classical estimation theory and Bayesian estimation theory. Then, Markov Random Fields are introduced. Their definition and their characteristic properties are presented. Finally the MAP-MRF framework is analyzed providing useful instruments (algorithms) to allow its implementation.

2.2 Labeling Problems

Many vision problem can be studied in terms of a labeling problem [1]. Formally speaking a labeling problem is specified defining a set of *sites* and a set of *labels*.

2. STATISTICAL SIGNAL PROCESSING

If we define a set of sites \mathcal{S} :

$$\mathcal{S} = \{1, \dots, S\} \quad (2.1)$$

and we define a set of label

$$\mathcal{L} = \{1, \dots, K\} \quad (2.2)$$

where S is the size of the site and K is the number of involved labels, a labeling problem consists in assigning a label f taken from the label set \mathcal{L} to a particular pixel s that belongs to \mathcal{S} . So a labeling can be seen as a mapping from \mathcal{S} to \mathcal{L} according to an established criterion. The set of labels:

$$F = \{f_1, \dots, f_S\} \quad (2.3)$$

is called a configuration of the set \mathcal{S} using the label set \mathcal{L} . The sites are normally the pixels of an image ($S = \text{row} \times \text{columns}$), while the labels are some particular features that have to be assigned to the involved pixels (i.e. pixels intensity in restoration problems, one of the possible mutually exclusive regions in segmentation problems).

When the S sites have the same label set, the configuration space, which is the set of all possible labels, is the given by:

$$\mathcal{F} = \mathcal{L} \times \mathcal{L} \times \mathcal{L} \dots \times \mathcal{L} = \mathcal{L}^S \quad (2.4)$$

Among this configuration set only a limited number of labels can be applied to a particular site, providing a good solution for the labeling problem. Due to the presence of many uncertainties in vision problems, i.e. noise and other degradation factors, the exact solution can be hardly determined. So the idea in labeling problems, is often to find an inexact but optimal configuration of the sites in terms of a specific criterion.

The best way to find the optimal solution for the labeling problem, due to the inevitable uncertainties in vision problems, is to exploit the instruments provided by the statistical estimation theory [2]. There exist two branches of the statistic estimation family. The first one is known as Classical Estimation, the second one is the known as Bayesian Estimation. This two branches give arise to two optimal criteria widely used: the Maximum Likelihood and the Maximum a Posteriori criteria, respectively [3].

2.3 Classical Statistical Estimation Theory

The aim of the statistical estimation theory is to estimate unknown parameters, considered as deterministic parameters, from a set of data, related to the unknown parameters via a known relationships [2]. Consider the label f the parameter to be estimated and \mathbf{d} the set of N available data.

$$\mathbf{d} = [d(0), d(1), \dots, d(N - 1)] \quad (2.5)$$

The target of the statistical estimation theory is to find an estimator \hat{f} for the unknown parameter f , starting from the available data \mathbf{d} and the data acquisition model. Since the collected data are implicitly random, they have to be described in probabilistic terms using the *probability density function* (pdf) $P(\mathbf{d}; f)$. Clearly, the data pdf depends on the parameters to be estimated. All the estimation techniques based on the pdf are called classical statistical methods [2]. Since the estimator \hat{f} is a function of the data

$$\hat{f} = g(\mathbf{d}) \quad (2.6)$$

it has to be considered as a random variable. Thus its assessment has to be performed in probabilistic terms (expectation and variance). With reference to the expectation, we look for an estimator that has to be unbiased. This means that its expectation has to be equal to the real value of the parameter:

$$E(\hat{f}) = f \quad (2.7)$$

In terms of variance, we look for an estimator that has the minimum possible variance. If this estimator can be found, it will be the so called *Minimum Variance Unbiased* (MVU) estimator. Note that the MVU estimator does not necessarily exist.

To evaluate the performances of any Unbiased estimator (not necessarily the MVU estimator) a fundamental instrument is given by the *Cramer Rao Lower Bounds* (CRLB) theorem. This theorem provides the lower bound for the variance of any unbiased estimator. Moreover, it provides a method to calculate under certain conditions, in the case it exists, the MVU estimator [2]

2. STATISTICAL SIGNAL PROCESSING

Theorem 1 *It is assumed that the data pdf $P(\mathbf{d}; f)$ satisfy the regularity condition:*

$$E \left(\frac{\partial \log P(\mathbf{d}; f)}{\partial f} \right) = 0 \quad \forall f \quad (2.8)$$

where the expectation is taken with respect to $P(\mathbf{d}; f)$. Then the variance of any unbiased estimator \hat{f} must satisfy the follow inequality:

$$\text{var}(\hat{f}) \geq \frac{1}{-E \left(\frac{\partial^2 \log P(\mathbf{d}; f)}{\partial f^2} \right)} \quad (2.9)$$

where the derivate is evaluated at the true value of f and the expectation is taken with respect to $p(\mathbf{d}; f)$. Moreover, an unbiased estimator that reaches the limits for any value of f can be found if and only if:

$$\left(\frac{\partial \log P(\mathbf{d}; f)}{\partial f} \right) = I(f) (g(\mathbf{d}) - f) \quad (2.10)$$

The estimator will be provided by:

$$\hat{f} = g(\mathbf{d}) \quad (2.11)$$

and the minimum variance will be:

$$\text{var}(\hat{f}) = \frac{1}{I(f)} \quad (2.12)$$

The term $I(f)$ is known as *Fisher Information* and it equal to:

$$I(f) = \frac{1}{-E \left(\frac{\partial^2 \log P(\mathbf{d}; f)}{\partial f^2} \right)} \quad (2.13)$$

Any unbiased estimator that reaches the CRLB is defined to be an *efficient* estimator.

Often the MVU estimator is difficult to be determined or in some cases it does not exist at all. This leads to the introduction of a class of estimator based on the so called *Likelihood* function. These estimators are easier to be determined and have a wide practical use.

When the data \mathbf{d} are observed, the pdf $P(\mathbf{d}; f)$ becomes a function of the unknown parameter. This function is called likelihood function. The Maximum

2.3 Classical Statistical Estimation Theory

Likelihood Estimator (MLE) is defined as the estimator that maximizes the Likelihood function.

$$\hat{f} = \arg \max_f P(\mathbf{d}; f) \quad (2.14)$$

The MLE can be shown to be asymptotically efficient. This means that, considering a large data set, the MLE is an unbiased estimator and it reaches the CRLB. Moreover, asymptotically it has a Gaussian distribution. Often, the asymptotic conditions are practically reached even if the data set is not so large.

Its practical definition and its asymptotic properties make the ML one of the most used optimization criterion in vision problem.

Referring to the labeling restoration problem, \mathbf{d} is the collection of noisy data referred to a pixel and f is the true intensity for the pixel that has to be estimated. Let consider, for example, the restoration problem in case of Additive White Gaussian Noise (AWGN) with a variance equal to σ^2 . Suppose we have N available data. The acquisition model is given by:

$$d(n) = f + w(n) \quad \text{for } n = 0 \dots N - 1 \quad (2.15)$$

The maximum Likelihood solution, obtained starting from the noise pdf (the joint pdf of all the noise samples $w(n)$), consist in

$$\hat{f} = \arg \max_f P(\mathbf{d}; f) = \arg \max_f \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (d(n) - f)^2\right) \quad (2.16)$$

which is equal to the the minimization of the Energy function

$$\hat{f} = \arg \min_f \sum_{n=0}^{N-1} (d(n) - f)^2 \quad (2.17)$$

In case of ill-posed problems (or ill-conditioned problems), to restore the capability to reach a solution, some constraints have to be added. These constraints are given by the *regularization* terms [4]. These regularization terms are based on physical hypothesis and force the algorithm to a solution with some specific characteristics. Considering the previous restoration example, if a regularization is needed, the solution can be obtained by minimizing an energy function given by:

$$\hat{f} = \arg \min_f \left[\sum_{n=0}^{N-1} (d(n) - f)^2 + \mu r(f) \right] \quad (2.18)$$

2. STATISTICAL SIGNAL PROCESSING

where $r(f)$ is the so called regularization term and μ is weighting factor which balances between the data constraints (Likelihood function) and regularization constraints.

2.4 Bayesian Estimation Theory

Bayesian Estimation theory is based on the hypothesis that the unknown parameter f is considered a random variable. With this assumption the configuration set of a labeling problem (2.3) becomes a *Random Field*. A random field

$$F = \{F_1, F_2, \dots, F_S\} \quad (2.19)$$

is a family of random variables defined on the set \mathcal{S} where each random variable F_s takes a value f_s in L . The joint event $(F_1 = f_1, F_2 = f_2, \dots, F_S = f_S)$, denoted with $F = f$, is a configuration of the field and it corresponds to a realization of the random field. A random field is described by its joint probability distribution $P_F(f)$ (prefix F is omitted in the following). The $P(f)$ is the a priori distribution of the labels.

According to the *Bayes* theory, when both Likelihood distribution $P(\mathbf{d}; f)$ and priori distribution $P(f)$ are known, the best solution is given by the *Bayesian Labeling*.

In Bayesian Labeling, the estimator is obtained by minimizing a Bayesian Risk \mathbf{R} . The Bayesian Risk \mathbf{R} is defined as:

$$\mathbf{R} = E(C(\hat{f}, f)) = \int_{-\infty}^{\infty} C(\hat{f}, f)P(f|\mathbf{d})df \quad (2.20)$$

where \mathbf{d} are the observed data, f is the true value, $C(\hat{f}, f)$ is a *Cost Function* and $P(f|\mathbf{d})$ is the posteriori distribution. The posteriori distribution is defined, according to Bayes Law as:

$$P(f|\mathbf{d}) = \frac{P(\mathbf{d}|f)P(f)}{P(\mathbf{d})} \quad (2.21)$$

where $P(\mathbf{d}|f)$ is the Likelihood function, $P(f)$ is the labels a priori distribution and $P(\mathbf{d})$ is the data distribution. The last term, $P(\mathbf{d})$, does not give any contribution to the estimation of the labels since it is a constant once the data are

observed (it does not depend on the labels). The choice of the Cost Function give arise to different estimators, with different characteristics. The choice of the Cost Function is driven by the different weight given to the difference between the estimate value \hat{f} and the true value f .

Mainly, two choices of the Cost Functions are made: a quadratic Cost Function and a “hit or miss” function (see fig. 2.1).

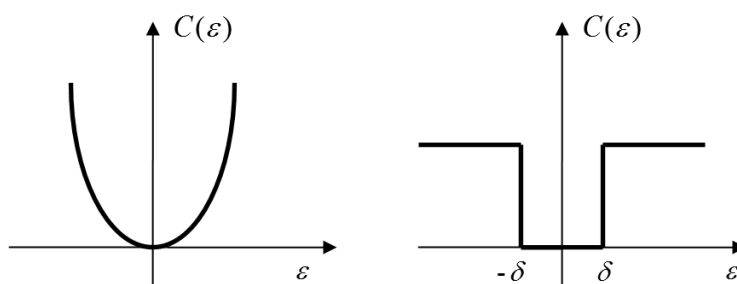


Figure 2.1: Cost Functions - Quadratic Cost Function and a “hit or miss” Cost Function where $\varepsilon = \|\hat{f} - f\|$

The quadratic Cost Function (figure 2.1) defined as:

$$C(\hat{f}, f) = \|\hat{f} - f\|^2 \tag{2.22}$$

assign a cost to the error between the estimate value \hat{f} and the true value f which grows according to a quadratic law.

The “hit or miss” function (figure 2.1)

$$C(\hat{f}, f) = \begin{cases} 0 & \text{if } \|\hat{f} - f\| \leq \delta \\ 1 & \text{otherwise} \end{cases} \tag{2.23}$$

assigns a null cost to small errors and a cost equal to 1 in case of errors bigger then a threshold δ . It can be shown that in case of a quadratic Cost Function (2.22), the minimization of the Bayesian Risk is equivalent to find an estimator which is the mean of the a posteriori distribution

$$\hat{f} = \int_{-\infty}^{\infty} f P(f|\mathbf{d}) df \tag{2.24}$$

2. STATISTICAL SIGNAL PROCESSING

In the case of the ‘hit or miss’ Cost Function (2.23) the minimization of the Bayesian Risk is equivalent to the maximization of the a posteriori distribution.

$$\hat{f} = \arg \max_f P(f|\mathbf{d}) = \arg \max_f \frac{P(\mathbf{d}|f)P(f)}{P(\mathbf{d})} \quad (2.25)$$

For a detailed demonstration see [1]. Because $P(\mathbf{d})$ is constant when the data are observed,

$$P(f|\mathbf{d}) \propto P(\mathbf{d}|f)P(f) \quad (2.26)$$

the MAP estimator is found by

$$\hat{f} = \arg \max_f P(f|\mathbf{d}) = \arg \max_f P(\mathbf{d}|f)P(f) \quad (2.27)$$

The latter estimator, known as the Maximum a Posteriori Estimator (MAP), is one of the most popular and used estimator in labeling problems.

Lets focus on the MAP estimator (2.27) and in particular lets consider the example of the restoration problem seen in the previous section. Suppose again that the collected pixels are corrupted by an AWGN (2.15). The a posteriori distribution is given by:

$$P(f|\mathbf{d}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp\left(-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (d(n) - f)^2\right) P(f) \quad (2.28)$$

The MAP estimator is given by the minimization of

$$\hat{f} = \arg \min_f \sum_{n=0}^{N-1} (d(n) - f)^2 + \lg P(f) \quad (2.29)$$

This expression is very similar to the expression (2.17) obtained in the ML case. There is a big and fundamental difference between (2.29) and (2.17). In the Bayesian case, the additive term $\lg P(f)$ does not have anything of arbitrary, but it comes directly from the labels distribution. In the Bayesian case the solution is not forced towards a particular solution. The obtained solution is the most natural possible solution.

As it can be easily understand, using the MAP approach, it is fundamental to well describe and define the a priori Information, in order to do not obtain optimal results. A common and widely spread method to model the a priori information

is the one based on Markov Random Field theory. This theory allows to model mainly any kind of processes, introducing contextual dependencies between the site. The MAP approach together with the MRF theory give arise to the MAP-MRF framework [5], which is one of the most used instruments in vision and imaging problems.

2.4.1 Markov Random Fields

Markov Random Field is a theory widely used in image processing. Even if it was already known before, it is with the work of Geman and Geman [5] in 1984 that MRF became very popular in image processing community. MRFs have been used in the last two decades in many vision problems, such as restoration, segmentation, edge detection and in reference to various application domains, such as satellite imaging, medical imaging.

In vision problem the introduction of contextual information is fundamental. Markov Random Field is a theory that provides an efficient instrument to model the a priori distribution $P(f)$ taking into account the contextual dependencies of physical phenomena. To introduce Markov Random Field is necessary to first introduce the concepts of *Neighborhood Systems* and *Cliques* (figure 2.2).

A *Neighborhood System* for the set \mathcal{S} is defined as

$$\mathcal{N} = \{N_s | \forall s \in \mathcal{S}\} \quad (2.30)$$

where N_s is the set of neighbors of the site s . The neighboring relationship has two properties:

- a site s is not neighbor to itself: $s \notin N_s; \forall s \in \mathcal{S}$;
- the relation is mutual: $s \in N_j \Leftrightarrow j \in N_s$;

Commonly two structures are used for N_s in case of a 2-dimensional lattice S : the 4-connexity neighboring system and the 8-connexity neighboring system (figure 2.2). In the first one each site has the top, bottom, right and left site as it neighbors. The latter includes also the 4 diagonal sites as neighbors.

A *Cliques* is a set of sites where each member of the set is a neighbor of all the other members. A clique can be of different order (figure 2.2). The set of

2. STATISTICAL SIGNAL PROCESSING

all the order cliques is denoted by $\mathbf{C} = C_1 \cup C_2 \cup \dots$. Normally, first-order and second order cliques are used, because higher order cliques cause too much computational complexity. First order and second order cliques over \mathcal{S} are denoted by $C_1 = \{s | s \in \mathcal{S}\}$ and $C_2 = \{s, j | s \in \mathcal{S} j \in N_s\}$. The interaction between sites in a clique is referred as *pixel interaction* and is modelled by a *clique potential*.

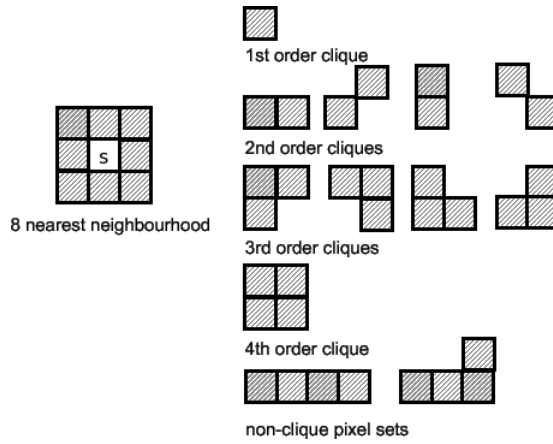


Figure 2.2: Neighborhood Systems and Cliques - 8-connectivity neighboring system and Cliques of different orders

Let now consider the random field defined in (2.19). A random field is a *Markov Random Field* (MRF) if and only if it satisfies two conditions:

$$P(f) > 0, \forall f \in \mathcal{F} \quad (2.31)$$

$$P(f_s | f_{S-\{s\}}) = P(f_s | f_{N_s}) \quad (2.32)$$

The first condition is related to the positivity of the joint distribution. The second condition, named Markovianity, implies that the information of one site s depends only on the information carried by its neighbors belonging to the considered neighboring system N_s . This second property allows to model the spatial interactions between sites.

An MRF can be specified either by the joint distribution or by the local conditional distributions. The first approach is preferred since the second one as

different disadvantages [6]. Thanks to equivalence between MRFs and *Gibbs* distributions established by the Hammersley-Clifford Theorem [7], an MRF can be conveniently analytically expressed in terms of joint distribution.

Theorem 2 *F is an MRF on a system \mathcal{S} with respect to the neighboring system \mathcal{N} if and only if F is a Gibbs Random Field on \mathcal{S} with respect to \mathcal{N} .*

Many proof of this theorem exist [1; 6]. This equivalence makes the MRFs a powerful and above all mathematically tractable instrument for signal processing. A Gibbs random field is specified by a Gibbs distribution:

$$P(f) = Z^{-1} \cdot \exp(-U(f)) = Z^{-1} \cdot \exp\left(-\sum_{C \in \mathbf{C}} V_C(f)\right) \quad (2.33)$$

where $U(f)$ is the *energy function* \mathbf{C} is the set of all cliques, Z is the so called *partition function* and $V_C(f)$ are the clique potentials. Thus, after specifying the energy function $U(f)$ or equivalently the clique potential functions $V_C(f)$, an MRF can be easily described using the same distribution as Gibbs Random Fields (2.33).

Therefore, the a priori information, can be modelled by an MRF using a Gibbs distribution, simply defining the energy function $U(f)$. The choice of the energy function is related to the particular problem to be solved.

Often the energy function $U(f)$ depends on some parameters θ . These parameters are introduced to allow the a priori model to be tuned or adjusted to achieve the best possible results for each available data set. This parameters are called *hyperparameters*. The hyperparameters θ have to be estimated (section 2.5.1).

2.4.1.1 Sampling an MRF

Once the MRF has been expressed in terms of Gibbs distribution, it is interesting and useful, in order to use the MRF, to see how to generate the samples of an MRF, starting from the Gibbs distribution $P(f)$ (2.33). To obtain realizations of a MRF two main algorithm have been developed: the *Gibbs Sampler* [5] and the *Metropolis Algorithm* [8].

Both of the methods are based on Markov Chain Monte Carlo (MCMC). In MCMC, a time sequence (Markov Chain) is generated, using the rule that at

2. STATISTICAL SIGNAL PROCESSING

each time t , the next random state $f^{(t+1)}$ is generated according to the Markov transition probability $P(f^{(t+1)}|f^{(t)})$. If an opportune transition probability is chosen, the chain converges after a certain time t_0 to the considered distribution $P(f)$. Thus the samples generated starting from the transition probability, after the time t_0 , can be considered as the samples of the $P(f)$. The time between the starting time and t_0 is known as *burn-in* period.

The steps of a MCMC are the follows:

1. set $t = 0$; generate an initial configuration f^0 ;
2. repeat until the maximum t is reached
 - generate the next configuration $f^{(t+1)}$ according to the Markov transition probability $P(f^{(t+1)}|f^{(t)})$
 - increase t
3. return the samples f^t , with $t > t_0$

So the procedure for the time $t < t_0$ does not provide samples, since the chain is not converged yet to the distribution $P(f)$. The useful samples are only the ones generated after the burn-in period. The choice of the transition probability function leads to different sampling algorithms, such as, for example, Gibbs sampler and Metropolis Algorithm.

The Gibbs sampler generates the next configuration based on the conditional probability. A candidate f_s' for the pixel s is randomly picked up from the conditional distribution $P(f_s'|f_{N_s})$. Basically the Gibbs sampler works as follows:

1. set an initial label f_s ;
2. repeat until a prescribed number of iterations have been performed (until the convergence is reached);
 - generate f_s' sampling from $P(f_s'|f_{N_s})$
 - set $f_s = f_s'$
3. return the samples f_s

The Gibbs sampler can be difficult to be implemented, since the conditional distribution has to be calculated, which can be a not easy task in some cases.

The Metropolis Algorithm generates the next configuration based on a energy change, rather than on conditional probability. At each step, for each site a label f' is generated starting from an uniform distribution in \mathcal{L} (\mathcal{L} is the set of the possible labels). The energy related to this new configuration is computed and it is compared to the energy related to the previous configuration. If the energy is decreased $\Delta U < 0$ (where ΔU is the difference between the two energies) the configuration change is accepted. If the energy is increased $\Delta U > 0$ the configuration change is accepted or refused with probability $p = \exp(-\Delta U)$ and $1 - p$. The process is continued until the equilibrium is reached. The Metropolis Algorithm generates the samples as follows:

1. set an initial label f ;
2. repeat until a prescribed number of iterations have been performed (until the convergence is reached);
 - generate the next configuration f' uniformly sampling from \mathcal{L} ;
 - calculate the energy difference $\Delta U = U(f') - U(f)$;
 - calculate $p = \min\{1, \exp(-\Delta U)\}$
 - generate a random number v between 0 and 1
 - if $v < p$ then set $f = f'$
3. return the samples f

It has been shown that both Gibbs sampler and Metropolis algorithm have the Gibbs distribution as equilibrium [5]. Comparatively speaking the Metropolis Algorithm is easier to be implemented since it does not involve the evaluation of the conditional distribution. Anyway, the choice of the sampling algorithm is strictly related to considered problem.

2.5 MAP-MRF framework

MRFs together with Bayes law give rise to the *MAP-MRF* framework introduced by Geman and Geman [5], which is one of the most popular framework used in labeling problems. In MAP-MRF framework the labeling is carried out using the Maximum a Posteriori (MAP) estimation technique using an a priori distribution modelled by an MRF.

The main aspect of this framework is that it considers the a posteriori distribution, under some hypothesis, as a MRF.

Let consider again the a posteriori distribution:

$$P(f|\mathbf{d}) = \frac{P(\mathbf{d}|f)P(f)}{P(\mathbf{d})} =_{\propto} P(\mathbf{d}|f)P(f) \quad (2.34)$$

where $P(f)$ is a MRF and the $P(\mathbf{d}|f)$ is the likelihood term that depends on the data. For the latter a valid hypothesis is to suppose the conditional independence of the sites:

$$P(\mathbf{d}|f) = \prod_s P(d_s|f_s) \quad (2.35)$$

Under this assumption and considering $P(f)$ equal to the first term of equation (2.33), the *a posteriori* distribution can be written as:

$$P(f|\mathbf{d}) =_{\propto} P(\mathbf{d}|f)P(f) =_{\propto} e^{\lg P(\mathbf{d}|f) - U(f)} =_{\propto} e^{-U(f|\mathbf{d})} \quad (2.36)$$

where

$$U(f|\mathbf{d}) = \sum_{s \in S} -\lg P(d_s|f_s) + \sum_{C \in \mathbf{C}} V_C(f) \quad (2.37)$$

Thus, the a posteriori distribution is still a Gibbs distribution. This means that the a posteriori distribution is a Markov Random Field.

The procedure of the MAP-MRF framework to solve the labeling problem can be summarized in three steps.

- Choose an appropriate MRF to model the a priori information: in order to perform this step a neighboring system, a set of cliques and cliques potential have to be defined.

- Calculate the a posteriori distribution adding the likelihood function (obtained from the acquisition model and the available data) to the a priori distribution.
- Estimate the hyperparameters of the a posteriori distribution.
- Find the MAP solution.

2.5.1 Parameter Estimation

As stated before, the MAP-MRF model can largely improve the quality of the achievable results in labeling problems. But, the MAP-MRF framework can also create potential error when the model is not well adapted to the data. To avoid these problems, and to allow the tuning of the model for different available data set, some hyperparameters θ are introduced in the distribution $P(f|\theta)$.

The hyperparameters are not a priori known, so they have to be estimated. To perform the hyperparameters estimation an ML estimation is carried out:

$$\hat{\theta} = \arg \max_{\theta} \lg P(f|\theta) \quad (2.38)$$

The previous estimation can be applied if the labels f are known. This is an ideal situation, because in this case it is possible to find the optimal value of the hyperparameter representing the image. Clearly, the hypothesis that the labels f are known is a nonsense, since the knowledge of the labels is what we are looking for. This case is the so called *complete-data* problem [1]. The *incomplete-data* problem [1], instead, is the estimation problem when the labels f are not known, so the hyperparameters have to be estimated indirectly from the available data \mathbf{d} (the classical situation).

2.5.1.1 Complete data problem

The complete data problem turns to be an interesting problem for two reasons. First, as it will be clear successively, the complete data problem is the basis for the solution of the incomplete data problem. It can be considered as a step for the incomplete data problem solution.

2. STATISTICAL SIGNAL PROCESSING

Secondly, the complete data problem can be used for simulation purposes, in order to assess the estimation quality in the incomplete data problem. Suppose, we have a known image. We can find the true hyperparameter using the complete data problem. Then we simulate the the direct problem, generate the measured data and estimate the hyperparameter using the incomplete data estimation on these data. The comparison between the complete data results and the incomplete data estimation provides an idea of estimation technique accuracy.

The estimation in the complete data case is carried out maximizing (2.38). Anyway, even in the complete data problem, finding the hyperparameter $\hat{\theta}$ can be a not easy task [9]. The maximization of (2.38) includes the evaluation of the partition function Z , which depends on the hyperparameter. A particular useful MRF model that allows to avoid the problems related to this maximization is the Generalized Gaussian MRF (GGMRF), for which a closed form of the estimation problem can be provided [10]. The distribution of a GGMRF model $P(f|\theta, p)$ is characterized by two hyperparameters θ and p , which are respectively a scale parameter and a shape parameter, with $p \in [1, 2]$. For our application, in the next chapters, we will use the shape parameters set to $p = 2$, providing the Gaussian MRF (GMRF).

The distribution of a GGMRF is given by:

$$\begin{aligned}
 P(f|\theta, p) &= \frac{1}{\theta^S z(p)} \cdot \exp\left(-\frac{1}{p}U(f|\theta, p)\right) = \\
 &= \frac{1}{\theta^S z(p)} \cdot \exp\left(-\frac{1}{p} \sum_{C \in \mathcal{C}} V_C(f|\theta, p)\right) = \\
 &= \frac{1}{\theta^S z(p)} \cdot \exp\left(-\frac{1}{\theta^p p} \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} (f_s - f_j)^p\right) \quad (2.39)
 \end{aligned}$$

where the partition function as been replaced by $Z(\theta) = \theta^S z(p)$ [10] and S are the number of involve sites.

For a given p , we perform the estimation of θ , using (2.38).

$$\hat{\theta}_{ML} = \arg \max_{\theta} \lg P(f|\theta, p) = \arg \max_{\theta} (-U(f|\theta, p) - S \log \theta - \log z(p)) \quad (2.40)$$

The ML estimator can be found setting:

$$\frac{\partial P(f|\theta, p)}{\partial \theta} \Big|_{\theta=\hat{\theta}_{ML}} = 0 \quad (2.41)$$

Using (2.41) and (2.40), we obtain:

$$\hat{\theta}_{ML} \frac{1}{S^p} \frac{\partial U(f|\theta, p)}{\partial \theta} \Big|_{\theta=\hat{\theta}_{ML}} = -1 \quad (2.42)$$

$U(f|\theta, p)$ can be written as $U(f|\theta, p) = \frac{1}{\theta^p} U(f|p)$ (scalability properties of GGMRF [10]). Thus

$$\hat{\theta}_{ML}^p = \frac{1}{S} U(f|p) = \frac{1}{S} \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} (f_s - f_j)^p \quad (2.43)$$

The estimation in closed form for the hyperparameter for a GGMRF, is a very interesting and useful result, that will be used in next chapters.

2.5.1.2 Incomplete data problem

The incomplete data problem can be effectively solved using the *Expectation-Maximization* (EM) algorithm [11]. The idea of EM algorithm is to first estimate the missing data f as given the current estimation of $\hat{\theta}$, use f to “complete” the data and then estimate a new $\hat{\theta}$ by maximizing the complete-data log likelihood. This procedure is done iteratively. However, the complete-data log likelihood is a function of the missing variables f . The EM algorithm proposes to use the expectation of the complete-data log likelihood to remove this dependency. The iterative steps of the EM algorithm are the follows:

- E-step: Compute the Expectation of the Conditional log-Likelihood function at the n -th step

$$Q(\theta|\theta^n) = E(\lg P(f|\theta) | \mathbf{d}, \theta^n) \quad (2.44)$$

- M-step: Maximize $Q(\theta|\theta^n)$ in order to update θ to the new value θ^{n+1}

$$\theta^{n+1} = \arg \max_{\theta} Q(\theta|\theta^n) \quad (2.45)$$

The E-step computes the conditional expectation of the unobservable labels f given the observed data \mathbf{d} and the current estimation θ^n . The M-step performs ML estimation as if there are no missing data.

2.5.2 Optimization in the MAP-MRF framework

In order to search the minimum energy configuration of a Markov Random Field (i.e. to maximize the a posteriori distribution) different algorithms have been developed. There exist deterministic and stochastic algorithms. Most of the deterministic algorithm works as follows: for a label f , the algorithm searches a new configuration f' in the neighborhood of that label f ; if f' leads to an improvement, $U(f') < U(f)$, than f is replaced by f' ; the process continues until no further improvement can be made.

In the stochastic search, the condition $U(f') < U(f)$ does not have to be necessarily satisfied. Whether to accept f' for which $U(f') > U(f)$ is decided to some probabilistic rules. Two of the most used minimization algorithms are *Simulated Annealing* [5] that belongs to the stochastic algorithms family and the *Iterated Conditional Modes* (ICM) [6] that belongs to deterministic one.

2.5.2.1 Simulated Annealing

To explain the algorithm a quick introduction to Gibbs distribution with temperature parameter is needed.

A Gibbs distribution with a temperature parameter T is defined as:

$$P_T(f) = Z(T)^{-1} \cdot \exp\left(\frac{-U(f)}{T}\right) \quad (2.46)$$

with $T > 0$.

It is interesting to see the different behavior of the distribution $P_T(f)$ at different values of the temperature T . It can be shown that:

- when $T \rightarrow \infty$ $P_T(f)$ converges to an uniform distribution, so all the configuration have the same probabilities;
- when $T \rightarrow 0$ $P_T(f)$ is uniformly distributed on the global minima (i.e. on the configuration providing the minimum energy).

The algorithm is based on the following steps:

1. set a starting temperature T very high

2. set an initial configuration f^0
3. at the n -th stage
 - simulate a configuration by sampling the Gibbs distribution (2.46) using the Metropolis Algorithm for a fixed temperature T^n (see section 2.4.1.1)
 - once the Metropolis algorithm converges to the equilibrium at current T^n , decrease T according to a carefully chosen schedule.
4. iterate until T is close to 0, when the system is “frozen” near the energy minimum.

The decreasing schedule normally used is

$$T^n > \frac{C}{\log(1+n)} \quad (2.47)$$

where C is a constant (normally set to $C = 3$).

Kirkpatrick et al.(1983) [12] choose another decreasing schedule given by:

$$T^n = kT^{(n-1)} \quad (2.48)$$

where k takes its values between 0.8 and 0.99.

Note that, opposite to the Gibbs sampler or the Metropolis algorithm that generate samples of the Gibbs distribution, the Simulated Annealing isolates only the most probable values (global minima of the energy).

Geman and Geman showed in [5] that Simulated Annealing can reach the global optimum regardless the initial configuration f^0 . Simulated Annealing allows to obtain the global optimum since it accepts to rise in energy. When the temperature decreases the risings become less possible. The temperature decreasing has to be performed very slowly to avoid the algorithm being trapped in a local minima.

The Simulated Annealing even though it can reach the global optimum, it can be excessively time demanding, since many configurations have to be simulated for each value of the temperature parameter T . Moreover the temperature parameter T has to be decreased very slowly in order to avoid local minima.

2. STATISTICAL SIGNAL PROCESSING

2.5.2.2 Iterated Conditional Modes

The Iterated Conditional Modes (ICM) algorithm has been proposed by Besag [6]. The ICM is a deterministic minimization algorithm. ICM is still an iterative algorithm that modifies at each step the labels f_s of each site, but differently from Simulated Annealing and others stochastic algorithm, the new value is chosen in a deterministic way.

The algorithm goes through all the sites s and for each site s :

- calculate the local conditional probabilities $P(f_s|\mathbf{d}, f_{N_s})$ for each possible value of f_s
- update the label assigned to the site s at the k -th iteration using the value that maximizes the conditional posteriori distribution:

$$f^{k+1} = \arg \max_{f_s} P(f_s|\mathbf{d}, f_{N_s}^k) \quad (2.49)$$

So, given the data \mathbf{d} and the other labels $f_{N_s}^k$, the algorithm sequentially updates each f_s^k into f_s^{k+1} by maximizing $P(f_s|\mathbf{d}, f_{N_s}^k)$ with respect to f_s . The process ends when the number of updates between two consecutive iteration becomes small.

This algorithm, opposite to the Simulated Annealing is very fast and is not computational heavy, since it just needs to evaluate the conditional a posteriori distributions. By the way the effectiveness of this algorithm is strictly related to the initialization f^0 provided to the algorithm. The ICM, as a matter of fact, can be seen as a Simulated Annealing but with a temperature parameter fixed to 0, $T = 0$. This doesn't make possible the energy rising of the Simulated Annealing, so the ICM can often be trapped in the local minimum nearest to the initial solution. How to properly set the initialization to obtain a good solution is not known. A natural choice for f^0 is the MLE.

2.6 Conclusions

In this chapter the labeling problem in computer vision has been faced. The labeling problem consists in finding the optimal configuration of a set of sites

(attribution of labels to sites, normally the pixels of an image). The labels are some particular features that have to be assigned to the involved sites. The configuration is searched according to an optimal criteria. Two main criteria are adopted in labeling problems: the Maximum Likelihood Estimation (MLE) and the Maximum a Posteriori (MAP) Estimation [1]. In both approaches the idea is to infer the estimation of the unknown parameter (the labels) from the available measured data and the acquisition model.

The MLE belongs to the Classical Estimation theory which is based on the Likelihood function (i.e. the probability density function of the data, once the data have been observed) which is a function of the unknown parameter. The unknown parameter is considered as a deterministic value. In the Classical Estimation theory a “good” estimator has to be unbiased (i.e the expectation of the estimator returns the true value of the parameter) and has to reach by the CRLB. It can be shown that the MLE is unbiased and reach the CRLB when enough data are available [2].

The MAP estimator belongs to the Bayesian Estimation theory which is based on Bayes’ Law. The main difference between the two theories is that in the Bayesian theory, the unknown parameter is considered as a random variable, which can be modelled by statistical distribution. A common way to model the a priori information is provided by Markov Random Fields (MRF). The estimator is provided by the maximization of the a posteriori distribution, given by the product of the Likelihood function and the a priori distribution.

Together the MAP approach and MRF give arise to the MAP-MRF framework introduced by Geman and Geman [5] which is the most used framework in labeling problems. The MAP-MRF framework can be summarized in 4 steps: chose an appropriate MRF model for the a priori information, calculate of the a posteriori distribution, estimate the parameters involved in the a posteriori distribution and finally find the MAP solution.

In order to estimate the involved parameters, the Expectation Maximization algorithm can be used [11]. In order to find the MAP solution, several optimization algorithms have been developed. Two of the most used in vision problem are the Simulated Annealing proposed by Geman and Geman [5] and the ICM proposed by Besag [6].

2. STATISTICAL SIGNAL PROCESSING

3

Multichannel SAR Interferometry

3.1 Introduction

In this chapter Multichannel InSAR systems are analysed. Multichannel InSAR systems consist in combining two or more independent interferograms of the same scene. Using the Multichannel InSAR configuration, together with a statistical-based approach, in order to efficiently combine the different acquisitions, it is possible to reconstruct the height profile of the considered scene. The height reconstruction of a scene can be a ill-posed problem. Two statistical approaches can be used: Maximum Likelihood and the Maximum a Posteriori approach. In both cases the uniqueness of the solution is assured. While in the first case a large number of interferograms is needed in order to achieve high reconstruction accuracy, using MAP approach together with Markov Random Field the number of interferograms needed is largely reduced.

In the first part of the chapter an introduction on Synthetic Aperture Radar is given. Then Interferometric configuration and the height reconstruction problem are introduced and deeply analysed. Finally the Multichannel InSAR configuration together with ML and MAP solution for the reconstruction problem are explained.

3. MULTICHANNEL SAR INTERFEROMETRY

3.2 Synthetic Aperture Radar

Synthetic Aperture Radar is a microwave active systems. Like others active systems, it emits microwave radiation to the ground and measures the electric field backscattered by the illuminated area. The aim of SAR processing is to transform the received signal into a high resolution image of the observed scene.

SAR, like the others Radar systems, is able to produce images from the backscattered signal, starting form the measurement of time delay between transmitted signal and received signal. The time delay is a function of the distance between the sensor and the scatterer.

The obtained SAR images are available in the two radar conventional coordinates, *azimuth* and *range*. The azimuth is the flight direction, while the range, which is orthogonal to the first one, is the looking direction of the antenna (i.e. the distance between the the sensor and the scatterer). Radar acquisition geometry is shown in figure 3.1.

With this geometry, a pixel on a SAR image is identified by an azimuth coordi-

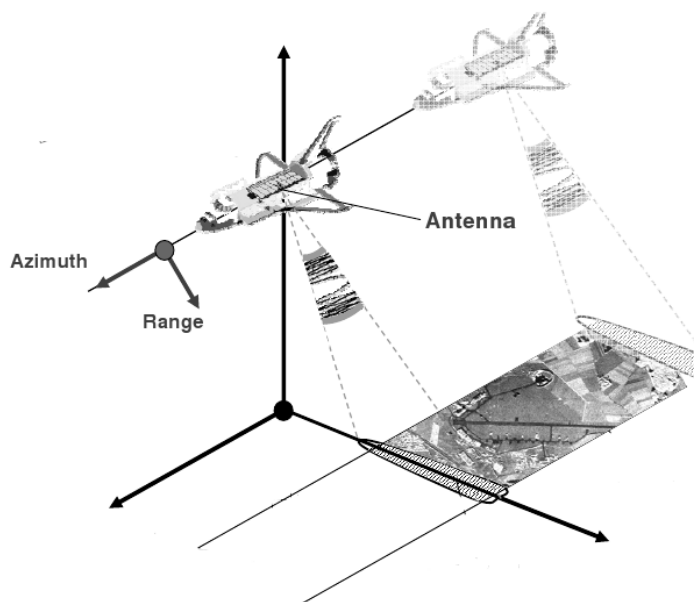


Figure 3.1: Radar acquisition geometry - The sensor flies in the azimuth direction and looks in the range direction

3.2 Synthetic Aperture Radar

nate and a range coordinate. SAR has different resolution on the two directions. The resolution is defined as the minimum distance that two scatterers must have in order to be solved. For the range direction the resolution is a function of the band of the compressed signal. If τ is the length of the compressed signal (and W is the bandwidth of the signal), the range achievable resolution is given by [13]:

$$\delta_r = \frac{c\tau}{2} = \frac{c}{2W} \quad (3.1)$$

where c is the speed of light. The range resolution is independent to the distance between the sensor and the scatterer points.

To achieve a high resolution, a linear frequency modulated signal is transmitted, the *chirp*. Using a chirp the energy of the signal is spread over a large bandwidth W . The energy is compressed again in a short time interval in the receiver using a matching filter.

The given definition of the range is the one referred to *slant range*. It exists, anyway, another definition of range: the *ground range*. Ground range is the direction orthogonal to the azimuth on the ground. Clearly, the resolution in ground range is different from the resolution in slant range. The relation between the two is given by:

$$\delta_g \simeq \frac{\delta_r}{\cos \psi} \quad (3.2)$$

where ψ is the so called grazing angle, which is defined as the angle between the radar line of sight and the local tangent plane at the point of the reflection on the earth. The ground range resolution is coarser than the slant range resolution. For azimuth resolution the SAR is different from conventional Radar systems. In conventional radar systems the azimuth resolution at a given range r_0 is given by [13]:

$$\delta_a \propto \frac{\lambda r_0}{L} \quad (3.3)$$

where λ is the wavelength and L is the antenna length. So in order to increase the azimuth resolution, it is necessary to use larger antennas and working at a small distance from the scatterers.

SAR systems take advantage of the fact that the response of a scatterer is contained in more than one single radar echo. A scatterer point, in fact, remains in the antenna beam for a significant period of time. So a scatter point is observed

3. MULTICHANNEL SAR INTERFEROMETRY

by the SAR from different positions during the movement of the antenna on its orbit (figure 3.2).

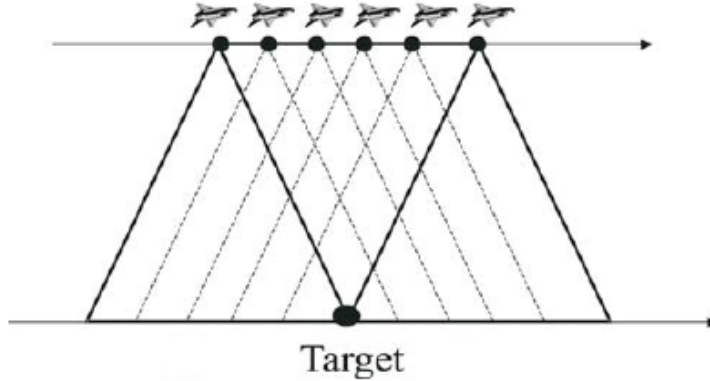


Figure 3.2: SAR acquisition system - The scatterer point is seen by the antenna from different positions

Exploiting this, instead of using a large antenna, by coherently combining the different echos relative to a scatterer point, SAR realizes a *synthetic enlarged antenna*, a sort of antenna array.

Using this configuration [13] the azimuth resolution is completely independent of the range distance and is determined only by the size of the real antenna L . At a given range r_0 SAR azimuth resolution is given by:

$$\delta_a = \frac{L}{2} \quad (3.4)$$

It would appear that a finer resolution can be obtained by simply making the real antenna length L very small. This is true, but some ambiguities conditions (in particular related to *Pulse Repetition Frequency* and the ground coverage capability of the SAR) place certain limits on this. A trade off between these and other involved parameters is necessary [14], [15].

3.2.1 SAR images

SAR images are not very easy to be interpreted, differently from optical remote sensing images. Due to the coherent data processing, SAR images show ampli-

tude values characterized by the *salt and pepper* effect and a phase randomly distributed. Inside a resolution cell, many different scatterers are located, providing a backscattered signal which is the coherent sum of all contributes characterized by a random distribution.. This coherent interference is the basis for the scintillation of SAR images, an effect often referred to as *speckle*. Thus, there is a wide variation in the SAR image, even over homogeneous areas [16].

Another interesting aspect of SAR images is due to the fact that since in SAR systems time delays are measured, points having the same distance to the antenna are located in the same position in the image. This can become a problem, or a source of misunderstanding of SAR images when areas with strong topography are investigated. In such cases, different geometrical distortion induced by the terrain can appear on SAR images. Three main geometrical distortion can appear: *foreshortening*, *layover* and *shadowing*. The three effects are shown in figure 3.3

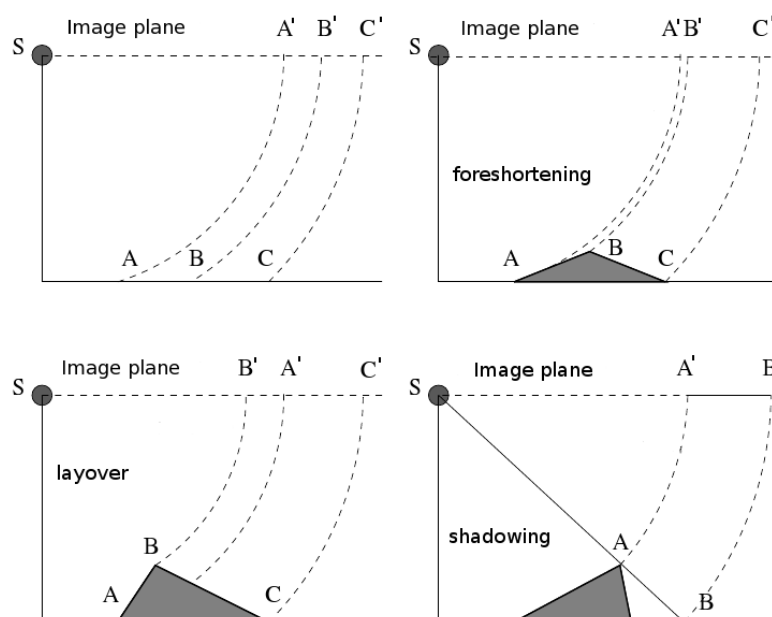


Figure 3.3: SAR geometrical distortions - Foreshortening, layover and shadowing effects in areas with strong topography

Let consider, first the foreshortening. A simplified terrain element (a triangular mountain) A, B, C , is illuminated by the radar. The echos from points A ,

3. MULTICHANNEL SAR INTERFEROMETRY

B , C are received at travel times that correspond in the slant range plane (image plane in figure 3.3) to A' B' and C' . It is clear that $\overline{A'B'}$ is much shorter than \overline{AB} providing a tilting of the top of the mountain towards the sensor.

The layover is illustrated in the third subfigure 3.3. The layover phenomenon arises when the slope of the mountain becomes bigger, from the side the slope is seen. The point on the top of the mountain B is nearer to the sensor, so it is mapped on the image plane before the point at the bottom of the mountain A (farther from the sensor). The echo from \overline{AB} is distributed in a reverse order on the image plane ($\overline{A'B'}$).

The shadowing effect is represented in the last subfigure of figure 3.3. This effect happens when the radar signal is shadowed by the topography. Backscattered information is lost and no signal is collected from point A to point B . This causes an area $\overline{A'B'}$ on the image plane that contains no signal, but only systems noise. The same geometrical distortions afflict urban scenario which is also characterized by strong topography (i.e. tall buildings). For a more detailed discussion and analysis on the geometrical distortion and in particular on the layover phenomenon in urban areas, refer to chapter 6.

3.2.2 SAR processing

A brief explanation of SAR processing is provided. For a more detailed one refer to [13].

In a SAR signal processor specific operations are required to convert the raw data (i.e. the acquired data) into an interpretable image. A SAR raw data is not an image yet since point targets are spread out in range and in azimuth. The spreading in range is due to the use of the chirp signal. The spreading in azimuth is due the fact that the echo of a point target is contained in many received radar pulse. The echo of a target point is received from the moving antenna for a time defined as *integration* time. The core of SAR processor is to correctly and coherently combine all this echo referred to the same target point received during the integration time. The focusing is needed both in azimuth and in range dimension, in order to create the image.

The focusing of the signal is made using some reference function, opportunely built. The reference functions are different for the range and for the azimuth compensation. Moreover for the azimuth compensation, the reference functions have to be adopted to the current range distance under investigation.

A simplified SAR processing sequence can be summarized as follows. The input is a complex signal (the raw data). After a one dimensional Fourier Transform (FFT) in range direction, each range line is multiplied with the FFT of the range reference function. After the inverse FFT the data are compressed in range. An FFT in azimuth direction is performed, followed by the multiplication with the FFT of the azimuth reference functions. An inverse FFT is applied, providing a compressed and focused SAR image. Note that after the range compression, the so called *range migration effect* can happen and has to be compensated [13].

3.3 SAR Interferometry

SAR image is a complex image, whose phase is randomly distributed (section 3.2.1). However, the phase signal contains a lot of useful information. *SAR Interferometric* systems are based on the use of the SAR complex image phase the . The term SAR Interferometry (InSAR) is referred to all methods that employ at least two complex SAR images, exploiting the phase of the complex SAR signals, in order to derive more information about an object respect to the information provided by a single SAR image [17]. To provide additive information it is necessary that at least one of the acquisition parameters is different from the first acquisition.

There exist two possible configuration of SAR Interferometry: *across track interferometry* and *along track interferometry*. In the across track configuration, 2 SAR sensors fly on two parallel flight lines and look at the ground from slightly different position. In the along track configuration the 2 sensors fly on the same flight direction, looking the scene from the same position but with a small temporal delay. This system is mainly used for measurement of ocean currents [18] and for moving target detection [19]. In this thesis the first configuration is considered.

3. MULTICHANNEL SAR INTERFEROMETRY

The geometry of an across track interferometric system is shown in figure 3.4. Two SAR systems (SAR1 and SAR2) look at the scene from two different positions. SAR1 is the so called *master antenna*, SAR2 is the *slave antenna*. The distance between the flight lines of the two SAR systems is called *baseline* B . The component of the baseline orthogonal to the look direction is called *orthogonal baseline* B_{\perp} . The distance between SAR1 and the target point and SAR2 and the target point are respectively R_1 and R_2 . $\theta_{1,2}$ denote the look angle of the sensors.

Across Track Interferometry is used to reconstruct earth topography, providing high precision Digital Elevation Model (DEM) of earth surface. For the across track configuration two approaches are possible. In the first one (*repeat pass interferometry*) the same sensor comes back and take a second image from a slightly different position. In the second approach two antennas are mounted on the sensor platform and operate in parallel (*single pass interferometry*).

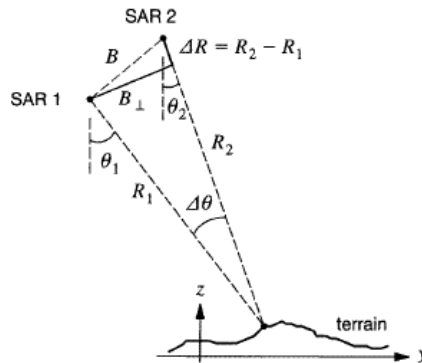


Figure 3.4: InSAR geometry - The scene is seen from two different positions

Let consider the two complex images collected by the two sensors SAR1 and SAR2, u_1 and u_2 . The interferogram ϕ of the two images is obtained by taking the phase of the product of the first image for the complex conjugate of the second one:

$$\phi = \angle u_1 u_2^* \quad (3.5)$$

$$\phi = \langle \phi_1 - \phi_2 \rangle_{2\pi} \quad (3.6)$$

where ϕ_1 and ϕ_2 are the phases of a the pixels in the complex images and $\langle . \rangle_{2\pi}$ is the modulo- 2π operation. It has to be noted, in fact, that the the acquisition system can only measure the wrapped value of the phase. The phase is measured in the principal interval $\phi \in [-\pi, \pi)$. The problem of recovering the absolute value of the phase will be addressed in section 3.3.3.

The phase ϕ_1 of a SAR image pixel consists of two terms (the same can be stated for ϕ_2): the first one is a geometrical term due to the distance between sensor and object (proportional to the range distance R_1). This is a strictly deterministic term. The second one is the *object-phase* which results from the scattering process on the object. The latter is a random value.

$$\phi_1 = \frac{4\pi}{\lambda} R_1 + \phi_{obj1} \quad (3.7)$$

$$\phi_2 = \frac{4\pi}{\lambda} R_2 + \phi_{obj2} \quad (3.8)$$

If the same object is observed two times from nearly the same position, it can be assumed that the object phase ϕ_{obj} is the same for both observations. In this case the phase difference between the two images (the interferogram) reflects the difference in the sensor-scatterer distance. It provides a very accurate measure of the difference in range [20], [21]:

$$\phi = \langle \frac{4\pi}{\lambda} \Delta R \rangle_{2\pi} \quad (3.9)$$

ΔR can be estimated with a very high precision. In fact, while the absolute values of range distances are only known with the sensor resolution (some meters), the interferometric phase can be measured with a subwavelength precision (millimeters).

Since there is a relation between ΔR and h , the height of the observed scene, this results allow to have very accurate and high resolution estimation of the height. Let see how to obtain the measure of the height (which is what we are interested in) from the knowledge of the phase.

3.3.1 Interferometer Height Sensibility

The measured phase is related to the height of that pixel. Let see how a topographic height change causes a proportional change in the interferometric phase.

3. MULTICHANNEL SAR INTERFEROMETRY

Figure 3.5 presents schematically the considered situation. The interferogram

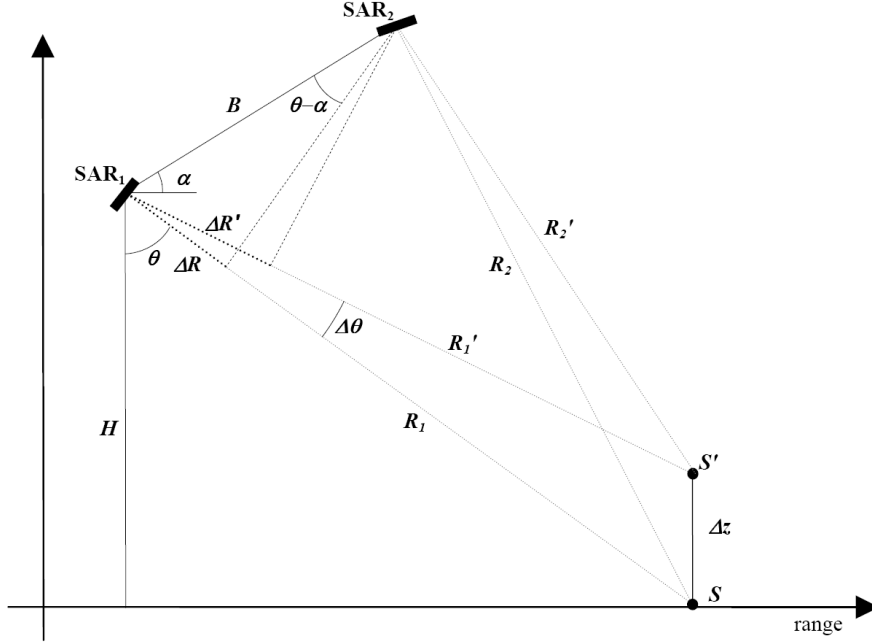


Figure 3.5: InSAR height Sensitivity - A topographic height change Δz causes a proportional change in the interferometric phase $\Delta\phi$

related to the point S is given by (3.9). ΔR , the difference between the two distances sensor-scatterer R_1 and R_2 can be written as the projection of the baseline B on the first antenna line of sight:

$$\Delta R \approx R_1 - R_2 = B \sin(\theta - \alpha) \quad (3.10)$$

where α is the angle between the baseline B and the horizontal plane. $B \sin(\theta - \alpha)$ is known as parallel baseline B_{\parallel} .

Considered point S' . It has a different height respect to point S . The difference in height of the two points is set to be Δz . The interferogram related to the point S' is given by (3.9):

$$\phi' = \left\langle \frac{4\pi}{\lambda} \Delta R' \right\rangle_{2\pi} \quad (3.11)$$

The variation Δz determines a variation of the look angle $\Delta\theta$. Thus $\Delta R'$ can be written as:

$$\Delta R' = R_1' - R_2' = B \sin(\theta + \Delta\theta - \alpha) \quad (3.12)$$

Let see, now the change in height Δz is reflected in the difference of the interferometric phases $\Delta\phi$:

$$\begin{aligned}\Delta\phi &= \langle \phi - \phi' \rangle_{2\pi} = \left\langle \frac{4\pi}{\lambda} (\Delta R - \Delta R') \right\rangle_{2\pi} = \\ &= \left\langle \frac{4\pi B}{\lambda} (\sin(\theta - \alpha) - \sin(\theta + \Delta\theta - \alpha)) \right\rangle_{2\pi} = \left\langle \frac{4\pi B}{\lambda} (\cos(\theta - \alpha)\Delta\theta) \right\rangle_{2\pi}\end{aligned}\tag{3.13}$$

where a small angle approximation has been used. $\Delta\theta$ is related to Δz according to the following relation:

$$\begin{aligned}\Delta z &= H - R_1 \cos \Delta\theta = R_1 \cos \theta - R_1 \cos(\theta + \Delta\theta) = \\ &= R_1 \cos \theta - R_1 \cos \theta \cos \Delta\theta + R_1 \sin \theta \sin \Delta\theta = \\ &= R_1 \sin \theta \Delta\theta\end{aligned}\tag{3.14}$$

where, again, a small angle approximation has been used. Substituting (3.14) in (3.13) we finally obtain:

$$\Delta\phi = \left\langle \frac{4\pi B}{\lambda R_1 \sin \theta} \cos(\theta - \alpha) \Delta z \right\rangle_{2\pi}\tag{3.15}$$

This relation provides the *Interferometer Height Sensibility* and provides the relation between the measured phase and the height in absence of noise (nominal phase).

3.3.2 Statistical Description

Let see how an interferogram can be characterized in statistical terms. As said before, the resolution cell contains many scatterers. A convenient way to model this situation can be obtained using a Gaussian scattering model [17]. This assumption allow to obtain analytical expressions for the statistical characterization of the interferograms [17].

Let consider the two SAR complex images u_1 and u_2 as two Jointly Circular Gaussian processes (both real part and imaginary part of each of the two processes are Gaussian). The joint probability density function is [22]:

$$P(\mathbf{w}) = \frac{1}{\pi^2 |\mathbf{C}|} \exp\left(-\mathbf{w}^{*T} \mathbf{C}^{-1} \mathbf{w}\right)\tag{3.16}$$

3. MULTICHANNEL SAR INTERFEROMETRY

where $\mathbf{w} = [u_1 u_2]^T$ and \mathbf{C} is the covariance matrix. the covariance matrix is defined as:

$$\mathbf{C} = \begin{pmatrix} E[|u_1|^2] & \gamma I \\ \gamma I & E[|u_2|^2] \end{pmatrix}$$

where $I = \sqrt{E|u_1|^2 E|u_2|^2}$ and γ is the so called *coherence coefficient*. The phase of the coherence coefficient is the expected interferometric phase of the considered pixel. Its magnitude of γ is related to the noise affecting the interferograms (see section 3.3.2.1 for more details on the coherence coefficient).

The marginal pdf of the phase can be derived from the joint pdf of phase and amplitude of the product $u_1 u_2^*$ [23].

Its expression is [22; 23]:

$$P(\phi) = \frac{1}{2\pi} \frac{1 - |\gamma|^2}{1 - |\gamma|^2 \cos(\phi - \phi^0)^2} \left(1 + \frac{|\gamma| \cos(\phi - \phi^0) \arccos(-|\gamma| \cos(\phi - \phi^0))}{(1 - |\gamma|^2 \cos(\phi - \phi^0)^2)^{1/2}} \right) \quad (3.17)$$

This pdf is described by two parameters ϕ^0 , the nominal phase used for the height reconstruction and $|\gamma|$ the magnitude of the coherence coefficient. If only an interval centered in ϕ^0 is considered, corresponding to the principal interval (i.e. $\phi \in [-\pi, \pi)$), the expectation and the variance of ϕ are given by [17]:

$$E[\phi] = \phi^0 \quad (3.18)$$

$$\sigma_\phi^2 = \frac{\pi^2}{3} - \pi \arcsin(|\gamma|) + \arcsin^2(|\gamma|) - \frac{Li_2(|\gamma|^2)}{2} \quad (3.19)$$

where $Li_2()$ is the Euler's dilogarithm. The behaviour of the phase pdf (3.17) for different values of the parameter $|\gamma|$ is shown in figure 3.6.

In case $|\gamma| = 0$ (noisy data), the pdf becomes a uniform pdf, while in case $|\gamma| = 1$ (noise-free data) it becomes a Dirac function.

3.3.2.1 Coherence Coefficient

The coherence coefficient γ is defined as the the normalized complex cross correlation between the two images [17] u_1 and u_2 :

$$\gamma = \frac{E[u_1 u_2^*]}{\sqrt{E|u_1|^2 E|u_2|^2}} \quad (3.20)$$

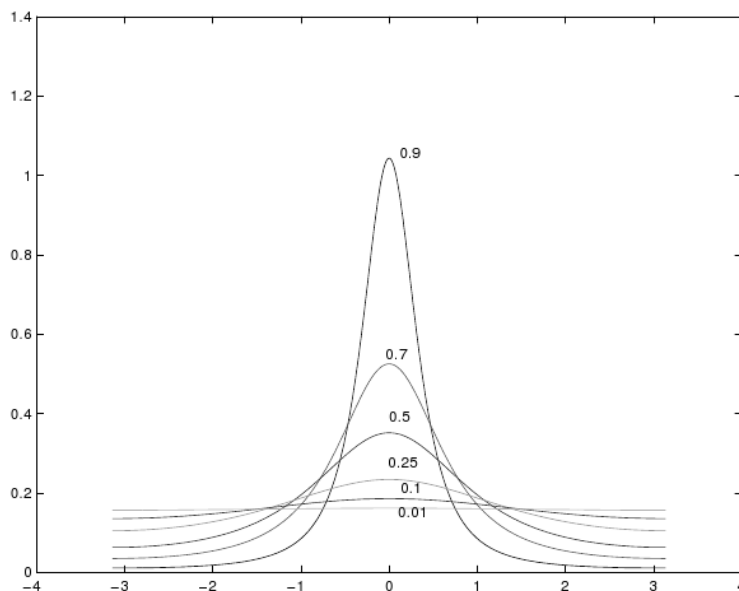


Figure 3.6: Interferometric phase pdf - Pdf of the interferometric phase (3.17), for different values of the coherence coefficient $|\gamma|$

An expectation is needed for its evaluation. Since only one measurement is available to calculate the coherence coefficient γ of a pixel, in practice the expectation value has to be replaced by a spatial averaging of the considered pixels with his neighboring pixels [17].

The magnitude of γ assumes value between 0 and 1 ($0 \leq |\gamma| \leq 1$). A value of $|\gamma| = 1$ comes from a perfect smooth phase, a value of $|\gamma| = 0$ means only noise. The interferometric phase is noisy because the assumption that the object phase ϕ_{obj} is identical in both SAR images is not always correct.

The coherence coefficient can be seen as the product of different terms:

$$\gamma = \gamma_{SNR}\gamma_t\gamma_H \quad (3.21)$$

These three terms influence the global γ behaviour.

γ_{SNR} is related to the influence of finite *Signal to Noise Ratio* (SNR). γ_t is related to the temporal decorrelation in repeat pass mode (see beginning of section 3.3). The decorrelation can happen because the scatterer orientation slightly changes in the two passes. Finally, γ_H describes the decorrelation caused by the fact that

3. MULTICHANNEL SAR INTERFEROMETRY

the two SAR images are acquired using different positions. This latter term can be written as:

$$|\gamma_H| = 1 - \frac{2 \cos \theta \delta_g B_\perp}{\lambda R_1} \quad (3.22)$$

where all the symbols have been defined previously. Equation (3.22) can be written as:

$$|\gamma_H| = 1 - \frac{B_\perp}{B_{crit}} \quad (3.23)$$

The B_{crit} is the so called *critical baseline*. It corresponds to a complete decorrelation of the interferometric images ($|\gamma_H| = 0$). The two interferometric images, in this case, cannot be used to form the interferogram (i.e. they see a different scene) [17].

3.3.3 Phase Unwrapping

SAR Interferometry allows to reconstruct height profile starting from the measure of the interferometric phase, using the relation (3.15). As mentioned before, the measured phase ϕ is known only in the principal interval $\phi \in [-\pi, \pi)$. The relation between height h and phase ϕ is, in absence of noise: has to written as:

$$\phi = \left\langle \frac{4\pi B}{\lambda R_1 \sin \theta} \cos(\theta - \alpha) h \right\rangle_{2\pi} \quad (3.24)$$

In order to reconstruct the height profile it is mandatory to restore the absolute values of the phase ($\varphi \in (-\infty, \infty)$). This problem is defined as *Phase Unwrapping* operation.

Suppose that:

$$\varphi = \phi + 2k\pi \quad (3.25)$$

where φ is the true value (the absolute value) and ϕ is the wrapped measured phase value (the modulo- 2π) and k is an integer accounting the number of 2π multiples. For a detailed treatment on phase unwrapping refer to [24]. Phase unwrapping (PhU) is the process of recovering φ from ϕ . This is, however, an ill-posed problem, since it admits an infinite number of solutions, if no further information is added. A regularization based on physical considerations, is needed to solve the problem.

Most phase unwrapping algorithms are effective only in case the *Itoh condition*

[25] is respected. The Itoh condition sets that the absolute value of phase difference between neighboring pixels is less than π . If this assumption is not violated, the absolute phase can be easily determined, up to a constant. Itoh condition can be violated in InSAR systems if the true phase surface is discontinuous, or if the wrapped phase measured phase is noisy. In both cases PhU operation becomes a very difficult task.

Phase unwrapping approaches belong mainly to three classes: *path following* [21], [26], *minimum norm* [27], [28], and statistical estimation approaches [29], [30].

Path following algorithms apply line integration schemes over the wrapped phase image, and basically rely on the assumption that Itoh condition holds along the integration path. Wherever this condition fails, different integration paths may lead to different unwrapped phase values. This approach includes the so-called *Branch Cuts* [24; 26]. In this approach the problem is not solved by a *global approach* (i.e all the observed phases are used to solve the problem), preventing these methods to be robust.

Minimum Norm methods try to find a phase solution ϕ for which the L^p norm of the difference between absolute phase differences and wrapped phase differences is minimized. This is, therefore, a global minimization in the sense that all the observed phases are used to compute the solution. In the case $p = 2$, we have a least squares method [31]. A drawback of the L^2 norm-based criterion is that it tends to smooth discontinuities, unless they are provided as binary weights. L^1 norm based criterion [28] performs better than L^2 norm in preserving discontinuities.

However, as said, these algorithms suffer of important limitation concerning discontinuities. Moreover most of these approaches do not exploit statistical properties of the noise present on the data so that these algorithms are not optimal from the information-theoretic point of view. The least squares method can provide the optimal solution only in the case of Gaussian noise. Since the interferometric noise is not Gaussian (the Gaussian nature of interferometric noise is an approximation) but it has a distribution provided by (3.17), the solution provided by least squares method is a suboptimal solution.

Statistical approaches relies on a data observation mechanism model. In case also the a priori knowledge of the phase is considered, these approaches are known as

3. MULTICHANNEL SAR INTERFEROMETRY

Bayesian approaches. Bayesian approaches can be optimal from the information-theoretic point of view but they are not able to restore uniqueness of the solution [29], [30].

An efficient solution for PhU problem, both from a information-theoretic point of view and from the ability to restore the uniqueness of the solution, even in case of high discontinuities, is the *Multichannel phase unwrapping* (MCPPhU). This method can be used in case several related SAR interferometric data are available [32], [33], [34], [35]. In the next section the MCPPhU solution will be presented and analysed.

3.4 Multichannel SAR Interferometry

In this section the Multichannel Phase Unwrapping (MCPPhU) approach is examined. MCPPhU approach consists in combining two or more interferograms of the same scene [32], [33], [36], [35]. The MCPPhU approach can be realized using two different configurations: *multifrequency* and *multibaseline* configurations.

Multifrequency configuration consists in using sensors operating at different frequencies and/or by partitioning the Fourier spectrum in non-overlapped sub-bands [36]. Since the sub-bands are not overlapped, it easy to show that the different obtained interferograms are statistically independent.

Suppose we have image pairs, each pair acquired in N non overlapped subbands (i.e. obtained by band pass filtering the full bandwidth images through N non overlapping Linear Filters). We combine the image pairs related to the same subbands in order to obtain the N interferograms:

$$\phi_n = \angle u_{1,n} u_{2,n}^* \text{ with } n = [1, 2, \dots, N] \quad (3.26)$$

where n is one of the N available subbands. The two images u_1 and u_2 , for each n , are assumed to be correlated narrowband Gaussian processes (section 3.3.2). The processes $u_{1,n}$ are Gaussian and uncorrelated from each other because they are obtained by passing the Gaussian process u_1 through N linear time invariant (LTI) filters whose passbands are non overlapped (figure 3.7). Thus, they are also statistically independent from each other. The same can be stated for $u_{2,m}$.

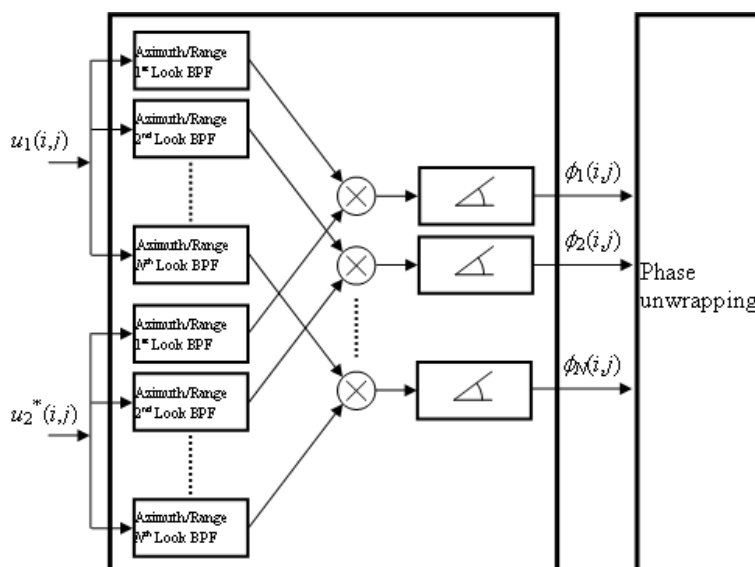


Figure 3.7: Multifrequency interferograms generation - Generation of multifrequency InSAR data using N linear time invariant filters

By the same considerations it follows that the processes $u_{1,n}$ $u_{2,m}$ are independent if $n \neq m$, while they remain correlated if $n = m$. Consequently, the N interferograms are statistically independent.

In order to have a larger number of subbands, and equivalently more independent interferograms, a *multilook* process can be applied. The multilook operation consists in dividing the azimuth bandwidth of the system in different non overlapped subbands. An example of multifrequency configuration using 2 different frequencies (no subbands on the range spectrum) and an azimuth spectrum partition (multilook) is shown in figure 3.8. Note that partition of the spectrum in several subbands, determines that the images are processed at a reduced bandwidth. Since the spatial resolution is proportional to the processing bandwidth, the final images exhibit a loss of spatial resolution [37]. So, if we reduce the bandwidth of a factor equal to k , the spatial resolution will results k times lower. A trade off between the spatial resolution and the number of subbands has to considered Multiple interferograms can be obtained also by considering the *multibaseline* configuration. The multibaseline approach consists in considering more than two SAR images obtained with slightly different view angles. The multibase-

3. MULTICHANNEL SAR INTERFEROMETRY

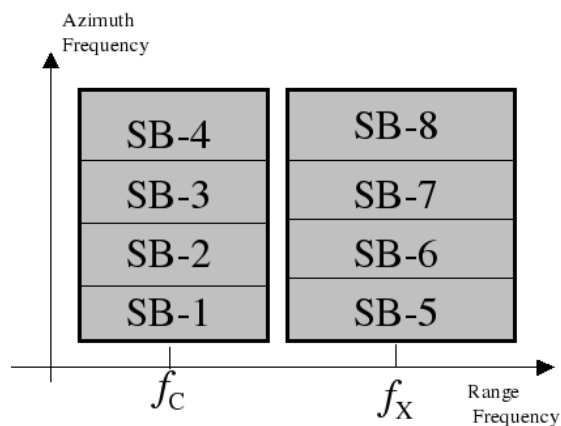


Figure 3.8: Multifrequency configuration - Fourier spectrum partition with 2 working frequencies (X -band and C -band) and 4 azimuth subbands, allowing $N = 8$ channels

line configuration in case of 2 baselines is shown in figure 3.9. The indepen-

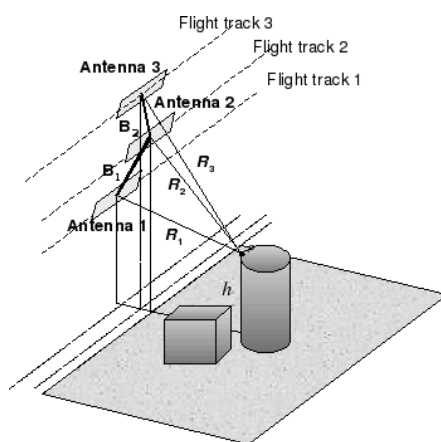


Figure 3.9: Multibaseline configuration - Dual baseline InSAR system geometry

dence between interferograms obtained in multibaseline configuration is not easily achieved. Consider a dual baseline configuration, the independence is satisfied only when four images are used to form the interferograms, and such four images are chosen in such a way that the spectrum of the first image pair and the spectrum of the second image pair are not overlapped. An interesting study on

multibaseline configuration, in case the interferograms are not statistical independent is provided by [38].

To achieve this condition, the two image pairs have to be acquired at a distance larger than the critical baseline (3.23) [36].

Starting from the relation (3.24) between the nominal phase ϕ^0 and the height of the considered pixel h , in case of N independent interferometric channels and considering the interferometric noise, the interferometric measured phase signal of a pixel can be written as:

$$\phi(s) = \langle \alpha_n h(s) + w_n(s) \rangle_{2\pi} \quad (3.27)$$

where

- $\alpha_n = \frac{4\pi B_{\perp,n}}{\lambda R_0 \sin(\theta)}$ in case of multibaseline method,
- $\alpha_n = \frac{4\pi B_{\perp}}{\lambda_n R_0 \sin(\theta)}$ in case of multifrequency method.

The index s refers to the pixel position inside the image of size S , index n to the considered channel which is one of the N possible interferograms (frequencies or baselines), w is the phase decorrelation noise. The others parameters have been previously defined (section 3.3). Defining these notations, the height reconstruction problem consists in estimating the height values $h(s)$ of the whole scene, using the $S \times N$ measured available wrapped phases $\phi_n(s)$.

Different methods to estimate the height values, exploiting N independent channels exist [32]. In the next sections the attention will be focused on the two methods proposed by [36; 39], both based on statistical signal processing. The first one relies on Maximum Likelihood (ML) estimation (section 2.3), the second one on Maximum a Posteriori (MAP) estimation (section 2.4).

3.4.1 Maximum Likelihood Estimation

In the Maximum Likelihood (ML) Estimation (section 2.3) approach the height value $h(s)$ is seen as a deterministic parameter. The approach is based on finding the value of $h(s)$ that maximizes the likelihood function. This problem can be seen as a labeling problem (section 2.2) where the sites are the image pixels s , the data are the measured interferograms $\phi(s)$ and the labels are the height value

3. MULTICHANNEL SAR INTERFEROMETRY

$h(s)$ to be assigned to the site s .

To explain how the methods works let start from the *singlechannel* case and then move to the *multichannel* one. Consider the probability density function of the interferometric measured phase $\phi(s)$ given the actual phase $\phi^0(s)$ (nominal phase) (3.17), in the principal interval ($\phi \in [-\pi, \pi)$) Substitute the nominal phase $\phi^0(s)$ with its relation to the height (3.15) and taking into account the periodicity of the cosine function we obtain the pdf of the measured phase $\phi(s)$ given the actual elevation $h(s)$:

$$P(\phi(s); h(s)) = \frac{1}{2\pi} \frac{1 - |\gamma(s)|^2}{1 - |\gamma(s)|^2 \cos(\phi(s) - \alpha h(s))^2} \times \left(1 + \frac{|\gamma(s)| \cos(\phi(s) - \alpha h(s)) \arccos(-|\gamma(s)| \cos(\phi(s) - \alpha h(s)))}{(1 - |\gamma(s)|^2 \cos(\phi(s) - \alpha h(s))^2)^{1/2}} \right) \quad (3.28)$$

Once the data have been observed (3.28) becomes the likelihood function, a function of the unknown parameter $h(s)$ (section 2.3). The plot of the likelihood function (3.28) for a given value of the interferometric phase $\phi(s)$ and for a fixed value of $\gamma(s)$ is shown in figure 3.10(a).

The likelihood function is periodic with period $T_h = \frac{2\pi}{\alpha} = \frac{R_1 \sin \theta \lambda}{B_\perp}$ and then it exhibits an infinite number of global maxima. Note that the period T_h depends on λ and on B_\perp , in particular it increases when λ increases (when working frequency decreases), and decreases when B_\perp increases. The ML solution is given by [36]:

$$\hat{h}(s)_{ML} = \arg \max_h P(\phi(s); h(s)) \quad (3.29)$$

that provides infinite solutions as likelihood function is periodic.

To avoid these ambiguities multichannel interferograms can be used. Considering N independent measured interferograms (ϕ_n con $n \in 1, 2, \dots, N$), obtained using N independent channels (section 3.4). The overall multifrequency likelihood function is given by the product of the N singlechannel likelihood functions (3.28):

$$F(\Phi(s); h(s)) = \prod_{n=1}^N P(\phi_n(s); h(s)) \quad (3.30)$$

where $\Phi(s) = [\phi_1(s), \phi_2(s), \dots, \phi_N(s)]^T$ is the measured wrapped data vector relative to the pixel s . The multiplication of N single-channel likelihood functions

3.4 Multichannel SAR Interferometry

with different periods (effect of different baselines and/or working frequencies) avoids multiple global maxima which are presented in single-channel likelihood function, at least in the range of interest for h . This is shown in figure 3.10 . The

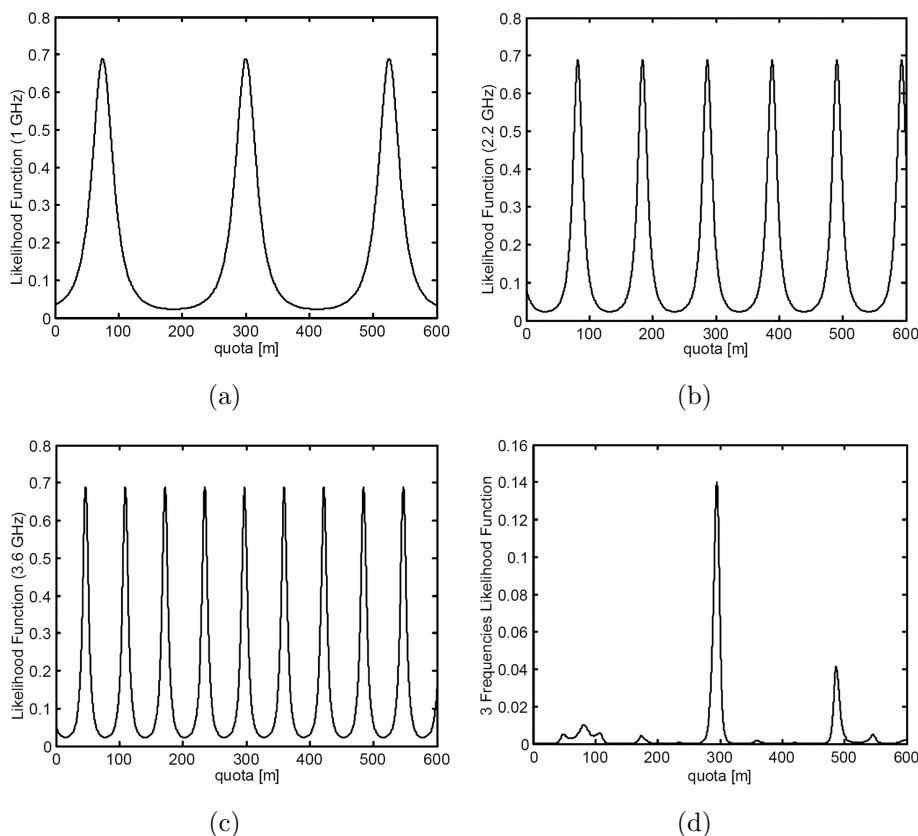


Figure 3.10: Singlechannel likelihood functions (3.10(a), 3.10(b), 3.10(c)) and multichannel likelihood function 3.10(d) given by the product of the previous three

plot of $N = 3$ singlechannel likelihood function, with different period, for given values of the interferometric phase ($\Phi(s) = [\phi_1(s), \phi_2(s), \phi_3(s)]^T$) are shown in figs 3.10(a),3.10(b), 3.10(c). The multichannel likelihood function (product of the $N = 3$ singlechannel likelihood function) is shown in figure 3.10(d). The multiplication of the $N = 3$ singlechannel likelihood function avoids multiple global maxima.

3. MULTICHANNEL SAR INTERFEROMETRY

The multichannel ML (MCML) solution is given by:

$$\hat{h}(s)_{MCML} = \arg \max_h F(\Phi(s); h(s)) = \arg \max_h \prod_{n=1}^N P(\phi_n(s); h(s)) \quad (3.31)$$

The solution of (3.30) represents the solution of the problem.

The MCML approach is able to restore the uniqueness of the solution. The drawback of MCML solution is that its performances depend significantly on the number of interferograms used. Reliable and good accuracy reconstructions (also in presence of critical scenario) can require a large number of channels, especially in presence of low coherence areas [36]. An effective way to reduce the number of needed interferograms and to increase the reconstruction accuracy is the MAP method proposed by [39].

3.4.2 Maximum a Posteriori Estimation

In the Maximum a Posteriori (MAP) Estimation approach (section 2.3) the height value $h(s)$ is seen as a random variable, modelled by an a priori distribution. Therefore, introducing an a priori information about the profile, the MAP approach is based on finding the value of $h(s)$ that maximizes the a posteriori distribution [39]. Like the ML estimation, also this problem can be seen as a labeling problem (section 2.2) where the sites are the image pixels s , the data are the measured interferograms $\phi(s)$ and the labels are the height value $h(s)$ to be assigned to the site s . In this case the labels are seen as random variables.

Let $\mathbf{h} = [h(1)h(2) \cdots h(S)]^T$ the unknown image, seen as a realization of the random vector \mathbf{H} . The unknown random image can be modelled by some statistical distribution. The distribution is usually defined in such a way to assign high probability to particular pixel configurations. In our case, being the unknown image representative of the elevation of a geographic area, a strong contextual pixel information is very likely to be. In section 2.4.1 we have seen that an excellent and powerful instrument to model the a priori information, taking into account the spatial constraints between image pixels are the Markov Random Fields (MRF).

Any MRF can be expressed in terms of Gibbs distribution (section 2.4.1, theorem

2). Therefore, the a priori can be modelled by (2.33):

$$P(\mathbf{h}, \boldsymbol{\theta}) = Z(\boldsymbol{\theta})^{-1} \cdot \exp(-U(\mathbf{h}, \boldsymbol{\theta})) \quad (3.32)$$

For our aim a valid choose for the energy function $U(\mathbf{h}, \boldsymbol{\theta})$ is the Local Gaussian MRF (LGMRF). According to this model, the energy function is set to be:

$$U(\mathbf{h}, \boldsymbol{\theta}) = \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} \frac{(h(s) - h(j))^2}{2\theta_{s,j}^2} \quad (3.33)$$

This model is actually a particular case of the GMRF (GGMRF (2.39) with shape parameter $p = 2$) [10]. Differently from the GMRF model the LGMRF is locally defined. In the GMRF the hyperparameter θ is a scalar (i.e. one hyperparameter for all the image). In the case of LGMRF the hyperparameter $\boldsymbol{\theta}$ is a vector (i.e. one hyperparameter for each couple of pixels). Note that the hyperparameter θ and the neighbouring system \mathcal{N}_s have been defined in (section 2.4.1). In the LGMRF model the hyperparameter $\theta_{s,j}$ is seen as an index of the spatial correlation of neighboring pixels. High value of $\theta_{s,j}$ means that the probability that to neighboring pixels s and j assume very different height values is high. Low value of $\theta_{s,j}$ means that the probability that $h(s)$ and $h(j)$ are different is small. By changing the hyperparameter $\theta_{s,j}$, the LGMRF can be adapted to describe the image local nature. This leads to a powerful and general model, well suited to represent a wide class of height profiles.

The local hyperparameters turns to be a very powerful instrument, because they allows to localize the flat regions and the discontinuities of the image, taking into account the local characteristics of the profile, and generating a hyperparameter map for the profile, very useful for the height reconstruction.

$\theta_{s,j}$ are not known but they have to be estimated (see section 2.5.1). Once the hyperparameter $\theta_{s,j}$ have been estimated the MAP solution is given by, following Bayes law

$$\begin{aligned} \hat{\mathbf{h}} &= \arg \max_{\mathbf{h}} F(\boldsymbol{\Phi}|\mathbf{h})P(\mathbf{h}, \boldsymbol{\theta}) \\ &= \arg \max_{\mathbf{h}} \log \left[\left(\prod_{s=1}^S F(\boldsymbol{\Phi}(s)|h(s)) \right) P(\mathbf{h}, \boldsymbol{\theta}) \right] \\ &= \arg \min_{\mathbf{h}} \left[-\sum_{s=1}^S \log F(\boldsymbol{\Phi}(s)|h(s)) + \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} \frac{(h(s) - h(j))^2}{2\theta_{s,j}^2} \right] \end{aligned} \quad (3.34)$$

3. MULTICHANNEL SAR INTERFEROMETRY

where

$$P_{post}(\mathbf{h}|\Phi) = F(\Phi|\mathbf{h})P(\mathbf{h}, \boldsymbol{\theta}) \quad (3.35)$$

is the a posteriori distribution of \mathbf{h} .

As said, the hyperparameter vector $\boldsymbol{\theta}$ has to be estimated (2.5.1). The estimation of $\boldsymbol{\theta}$ is carried out using Maximum Likelihood estimation.

Suppose the pixels height $h(s)$ are known. The estimation problem in this case is the *complete data* problem. An effective way to estimate the scalar hyperparameter θ in the complete data problem is provided in closed form by [10] (see section 2.5.1.1). Using the closed form (2.43) and considering a local hyperparameter and 8-connexity neighbouring system \mathcal{N}_s , the hyperparameter θ_s of the pixel s can be calculated in closed form as:

$$\hat{\theta}_s^2 = \sum_{j \in \mathcal{N}_s} \frac{(h(s) - h(j))^2}{9} \quad (3.36)$$

So if $h(s)$ is known for all the pixels image, (3.36) provides the estimation of the parameters $\hat{\theta}_s^2$ for the complete data problem.

Since $h(s)$ is not known (*incomplete data* problem), it has to be first estimated and then used to implement (3.33). To solve this problem an efficient algorithm is the Expectation Maximization algorithm, which is a technique for finding the MLE of hyperparameters in case of incomplete data problems (section 2.5.1). Using the EM algorithm the hyperparameter estimation at the $(k + 1)$ -th iteration is

$$\hat{\theta}_s^2(k + 1) = E \left(\sum_{j \in \mathcal{N}_s} \frac{(h(s) - h(j))^2}{9} \middle| \Phi, \hat{\theta}_s^2(k) \right) \quad (3.37)$$

For the evaluation of the conditional expectation, thanks to the ergodicity of the a posteriori distribution [5], the expectation can be approximated by sample averaging. Samples of $h(s)$ at the k -th iteration are mandatory. The sampling is performed starting from the a posteriori distribution of \mathbf{h} (3.35). The a posteriori distribution can be considered as the probability density function of a MRF. It has the expression of a Gibbs distribution. So, to obtain realizations of this MRF the *Gibbs Sampler* [5] or the *Metropolis Algorithm* [8] can be used (see 2.4.1.1). Once estimated all the hyperparameters, the hyperparameter $\hat{\theta}_{p,q}$ is set to be the mean of the hyperparameter $\hat{\theta}_s$ and of the hyperparameter $\hat{\theta}_j$.

After the hyperparameter estimation is performed, the height reconstruction is carried out using (3.34). In order to isolate the modal value of (3.34), the Simulated Annealing (2.5.2.1) can be used. This can be excessively time demanding, so a *semi-deterministic* solution is achieved using the ICM (section 2.5.2.2) algorithm. The ICM is initialized with high-probability samples of the image generated in the hyperparameter estimation step (3.37). This approach guarantees a good grade of effectiveness in terms of convergence, and at the same time a realistic convergence time.

The presented MAP method can be particularly useful to solve the reconstruction problem (3.24) when no *permanent scatterers* [40] are present in the considered area (coherent target reflecting structures on the ground) and when the interferograms are characterized by low coherence coefficients.

Starting from the point that the multichannel methods (both the MAP and ML) are able, exploiting different channels, to restore the uniqueness of the solution for the height reconstruction problem (section 3.24) and to get the correct solution, an interesting work on the high accuracy achievable using the Multichannel (in particular a multibaseline configuration) approaches has been recently conducted [41]. The results of this analysis show that MAP approach always outperforms the ML one and its attainable accuracy is not sensitive to the baseline choice, while mainly depends on the ground slope.

3.5 Conclusions

In this chapter the Multichannel Interferometric Synthetic Aperture Radar systems have been considered. Synthetic Aperture Radar are microwave active systems that like others active systems, emit microwave radiation to the ground and measure the electric field backscattered by the illuminated area [37]. Differently from other radar systems, SAR can achieve high azimuth resolution, even if mounted on moving platforms far away from the target (satellites).

InSAR systems consist in combining at least two complex SAR images of the same scene, seen from two slightly different position. Exploiting the phase of the two complex images, InSAR systems are able to reconstruct the height profile of

3. MULTICHANNEL SAR INTERFEROMETRY

earth surfaces. In order to perform the height reconstruction the Phase Unwrapping (PhU) operation has to be performed [17].

The PhU operation consists in the transformation of the wrapped interferometric phase signal (that assumes its values in the principal interval), into an unwrapped phase signal whose values are not limited to the principal interval. As the unwrapped phase signal is related to the height profile of the earth surface, the PhU operation allows to recover the terrain topographic map.

The PhU problem is an ill-posed problem, since infinite height profiles exist such that they can be wrapped in the same interferogram, so this problem must be regularized. Many techniques have been applied trying to regularize the PhU problem [24], [28]. Often these techniques are not able to restore the uniqueness of the solution when the Itoh condition is not satisfied (images characterized by high discontinuities between adjacent pixels, bigger than π). Moreover, most of them do not exploit statistical properties of the noise present on the data, so that these algorithms are not optimal from the information theoretic point of view.

To recover the solution uniqueness without imposing any constraint on the height profile of the ground surface even in the case where the Itoh condition is violated, different solutions have been designed exploiting the availability of many interferograms referred to the same scene, obtained, for example, using different baselines or different frequencies. This case is generally indicated as multichannel InSAR configuration [32], [33], [36]. In such situation, to solve the problem of the height reconstruction, a solution consists of adopting a statistical-based approach, in order to efficiently combine the different acquisitions. Two main approaches have been developed: a Maximum Likelihood estimation approach and a Maximum a Posteriori (MAP) estimation one. Both of the approaches are able to restore the uniqueness of the solution. The drawback of ML approach is the large number of interferograms required to assure high reconstruction accuracy [33], [36].

This problem can be overcome using the Maximum a Posteriori (MAP) estimation approach, where the number of needed interferograms in order to obtain high reconstruction accuracy, is reduced respect to the ML case. This is possible thanks to the introduction of an a priori statistical term to model the unknown height profile. In particular a Gaussian Markov Random Field is used, characterized by a set of local hyperparameters (Local Gaussian MRF - LGMRF). By

3.5 Conclusions

changing the hyperparameters the LGMRF can be adapted to describe the image local nature. This leads to a powerful and general model, well suited to represent a wide class of height profiles.

3. MULTICHANNEL SAR INTERFEROMETRY

4

Phase Offset Estimation in Multichannel SAR Interferometry

4.1 Introduction

In this chapter the problem of the presence of *phase offset* in the real interferometric data is considered. The interferograms are known except for a not known slowly varying phase. This phase is the results of several factors. Estimation of the offsets is mandatory in order to implement the Multichannel phase unwrapping technique and retrieve the absolute height profile from data, whatever statistical estimation method (ML or MAP) we adopt. The estimation of the phase offsets is based on Maximum Likelihood estimation, using the Expectation Maximization. The procedure is validated and assessed using simulations, and then it is applied to a real data case. In the first part of the chapter the phase offset problem is exposed. Then the estimation procedure of the phase offsets is considered, and finally some results both on simulated and real data are shown in order to prove the effectiveness of the method.

4.2 Phase Offset in Multichannel InSAR Systems

InSAR systems allow to estimate the altimetric profile of the Earth Surface using the PhU operation (section 3.3). PhU is an ill posed problem, admitting an infinite number of solutions. Using the multichannel InSAR configuration (section 3.4), it is possible to restore the uniqueness of the solution. In order to efficiently combine the different channels acquisitions, a statistical based approach is used (Maximum Likelihood and Maximum a Posteriori). For more details see sections 3.4.1 and 3.4.2.

However, despite the particular solution method used, the problem with multichannel Interferometry is that interferograms can be affected from the presence of undetermined phase offsets. The multichannel interferograms are known except for a not known slowly varying phase, which makes difficult to get correct height estimation in any case. This phase offset is the combination of different factors. In particular it results from the combination of a non-perfect synchronization in the acquisition system, from SAR processing errors, from some atmospheric effects and from the parallel baselines uncertainty [42].

The phase offset is different for each different interferogram, but it can be reasonably supposed to be slowly varying or spatially homogeneous on a single interferogram, for not very large acquisition areas. Estimation of these offsets is mandatory in order to retrieve the absolute height profile from data, whatever statistical estimation method (ML or MAP) we adopt. Considering that each interferogram has a different offset value, the problem can be formulated as the estimation of a set of unknown real values, starting from the knowledge of the interferograms, the acquisition model, and the noise statistics. Once the phase offsets values have been estimated, the multichannel height reconstruction ML (3.31) or MAP (3.34) procedure can be applied to retrieve the height profile.

Then, if we suppose that each interferogram is known except for a phase offset, generally speaking we have:

$$\phi' = \phi + \psi \tag{4.1}$$

where ϕ is the “normal” noisy wrapped phase in a pixel, while ψ is the phase offset. We suppose the phase offset is constant, in the area where the height estimation is performed, for a fixed baseline (or frequency); then, we have to estimate a vector of constant phases:

$$\boldsymbol{\psi} = [\psi_1, \psi_2, \dots, \psi_N]^T \quad (4.2)$$

where ψ_n is the phase offset of the N -th interferogram, associated to the N -th channel. To restore the correct relation between the interferograms, and consequently to make possible to perform the height’s reconstruction, an estimation of these offsets is mandatory.

4.3 Phase Offset Estimation

Let us now consider a subset of the entire image pixels S . In particular M pixels are chosen. The adopted criterion leads to consider the highest coherence pixels among the image. Only the data related to such subset are used to estimate phase offsets. Let introduce the vector

$$\mathbf{s} = [s(1)s(2) \dots s(M)]^T \quad (4.3)$$

with $s(m) \in \{0, \dots, S\}$, which is the collection of the indexes that identify the considered M pixels in the image. Then, the collected phase data values can be written as:

$$\boldsymbol{\phi}' = \{\phi'_{s(m)n}\} \quad (4.4)$$

whit $m = 1, \dots, M$ and $n = 1, \dots, N$. M is the number of considered pixels and N is the number of available interferograms.

The collected phase data values are related to the unknown offset through (4.1), which can be written, considering only M high coherence pixels, as:

$$\phi'_{s(m)n} = \phi_{s(m)n} + \psi_n \quad (4.5)$$

This is shown in figure 4.1.

Moreover, the vector collecting the height values, corresponding to the considered

4. PHASE OFFSET ESTIMATION

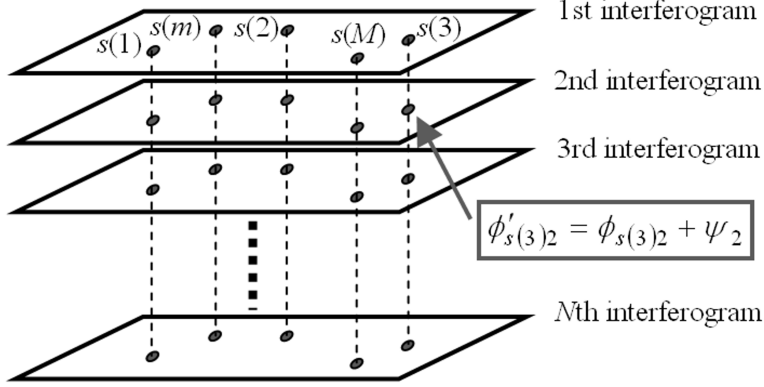


Figure 4.1: High Coherence Pixels - Choice of the M highest coherence pixels in the N available interferograms

cluster of pixels in the image, can be expressed as:

$$\mathbf{h}_M = [h_{s(1)} h_{s(2)} \cdots h_{s(M)}]^T \quad (4.6)$$

where $h_{s(m)}$ is the m -th pixel height value in the considered set. The value assumed by this vector has to be estimated, too. So, the data of the problem are $M \times N$ (i.e. the elements of vector ϕ'), and the unknowns of the problem are $M + N$ (i.e. the elements of vectors \mathbf{h}_M and ψ).

This problem is solved using a *complete data* and an *incomplete data* procedure (see section 2.5.1). First the problem of estimating the offset values is solved supposing that height values are known (i.e. complete data problem). Then, the incomplete data problem is solved, where the only available data are the measured phase values (4.4). Let start with the complete data problem. To estimate the vector ψ , the Maximum Likelihood (ML) estimation method is adopted. In order to use this approach, first the likelihood function of our problem has to be built. In particular, exploiting the statistical independence between the phase values at different acquisitions (frequency/baseline), and in different pixels, we can factorize the likelihood function, as the product of each single likelihood function. The single likelihood function is the likelihood function for a single pixel and for a single frequency/baseline. The Likelihood function of our problem is:

$$L(\phi', h_M | \psi) = \prod_{n=1}^N \prod_{m=1}^M P(\phi'_{s(m)n}, h_{s(m)} | \psi_n) \quad (4.7)$$

where $P(\phi'_{s(m)n}, h_{s(m)} | \psi_n)$ is the single likelihood function. The single likelihood function can be obtained starting from the likelihood function (3.28) adding the unknown phase offsets:

$$\begin{aligned}
 P(\phi'_{s(m)n}, h_{s(m)} | \psi_n) &= \frac{1}{2\pi} \frac{1 - |\gamma_{s(m)n}|^2}{1 - |\gamma_{s(m)n}|^2 \cos(r_{s(m)n})^2} \\
 &\times \left(1 + \frac{|\gamma_{s(m)n}| \cos(r_{s(m)n)} \arccos[-|\gamma_{s(m)n}| \cos(r_{s(m)n})]}{(1 - |\gamma_{s(m)n}|^2 \cos(r_{s(m)n})^2)^{1/2}} \right)
 \end{aligned} \tag{4.8}$$

where:

$$r_{s(m)n} = \phi'_{s(m)n} - \alpha_n h_{s(m)} - \psi_n \tag{4.9}$$

where $\gamma_{s(m)n}$ is the coherence at pixel $s(m)$ in the n -th interferogram and α_n has been defined in case of multibaseline and in case of multifrequency in section 3.4. The ML estimation consists of the maximization, respect to the vector $\boldsymbol{\psi}$ of the multichannel likelihood function (4.7). The maximum of the likelihood function can be found in an exhaustive way, supposing that the true height values are known (complete data case). Moreover, the ML estimation can be decoupled in N different maximization problems:

$$\begin{aligned}
 \hat{\psi}_n &= \arg \max_{\psi_n} L(\boldsymbol{\phi}', \mathbf{h}_M | \boldsymbol{\psi}) = \\
 &= \arg \max_{\psi_n} \prod_{n=1}^N \prod_{m=1}^M P(\phi'_{s(m)n}, h_{s(m)} | \psi_n)
 \end{aligned} \tag{4.10}$$

For simplicity in the notation let set:

$$U_n(\mathbf{h}_M) = \arg \max_{\psi_n} \prod_{n=1}^N \prod_{m=1}^M P(\phi'_{s(m)n}, h_{s(m)} | \psi_n) \tag{4.11}$$

To solve the problem in the incomplete data case, *Expectation Maximization* (EM) algorithm is used, which is an iterative algorithm, based on alternating a maximization step to the evaluation of an expected value (see section 2.5.1). This method exploits the complete-data solution in (4.11), and converges to a local maxima of the likelihood function, iteratively updating the parameter value at $(i+1)$ -th iteration, by evaluating the following expectation:

$$\hat{\psi}_n(i+1) = E[U_n(\mathbf{H}_M) | \boldsymbol{\phi}', \boldsymbol{\psi}(i)] \tag{4.12}$$

4. PHASE OFFSET ESTIMATION

with $n = 1, \dots, N$. \mathbf{H}_M is a vector of random variables. For the evaluation of the expected value over \mathbf{H}_M an analytically closed form can not be derived in this case. However, thanks to the ergodicity of the a-posteriori distribution [5], this expected value can be approximated by “time” averaging. Then, for the latter it is necessary to sample from the a posteriori distribution at the i -th step, whose expression is:

$$P_{post}^i(\mathbf{h}_M | \phi', \psi = \psi(i)) \propto L(\phi', \psi = \psi(i) | \mathbf{h}_M) P(\mathbf{h}_M) \quad (4.13)$$

For the a priori distribution $P(\mathbf{h}_M)$, it is supposed, regarding the offset estimation, that a uniform distribution is adopted. This is done to leave the estimation procedure complexity as lower as possible. In order to generate the samples needed to realize the expectation step, we adopt the Metropolis Algorithm; for more details about its implementation, see section 2.4.1.1.

4.4 Simulated Experiments

In order to assess the validity of the offset estimation procedure (section 4.3), in this section numerical example for the phase offset estimation are presented. As first example, let consider the height profile shown in figure 4.2. It consists of a 64×64 pixels linear ramp with a 1.1 m/pixels slope. 5 interferometric phases obtained with 5 different baselines have been simulated. The simulated system parameters are from the ENVISAT-ASAR sensor. The adopted baseline set is inspired to a real data set acquired by the ENVISAT-ASAR sensor, and is the following: [237.72 255.90 322.59 399.92 544.81 m]. For the simulation of the interferograms the coherence value is assumed homogeneous ($\gamma = 0.85$). The vector of phase offsets is assumed equal $\psi = [2.3 \ 2 \ 2.5 \ 0.8 \ 1]^T$. To perform the phase estimation $M = 900$ pixels have been used.

The results of the phase offset estimation are shown in figure 4.3(a) and in the third row of table 4.1. Once the phase offsets are estimated, $\hat{\psi}$ are used for the MAP multichannel reconstruction of the profile. A cut of the height profile (solid line) together with a cut of the reconstructed height profile obtained with the estimated phase offsets (dotted line) and without the phase offset estimation (dashed line) are shown in figure 4.3(b). The normalized reconstruction square

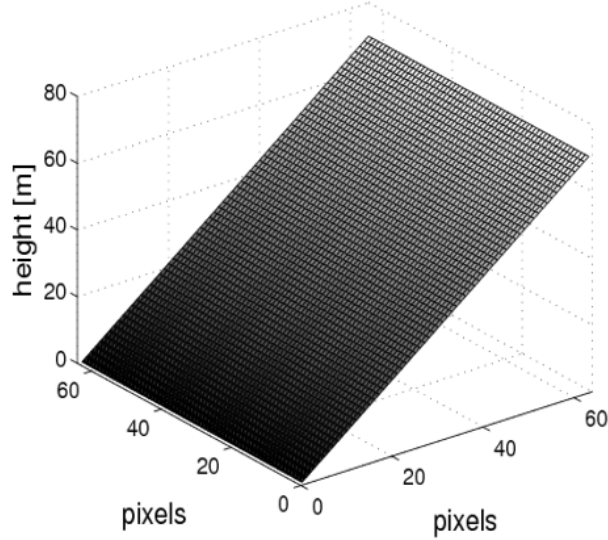


Figure 4.2: Phase offset simulated experiment - ramp profile - Reference height profile

Interferogram	1st	2nd	3rd	4th	5th
True Offsets	2.3	2	2.5	1	0.8
Estimated Offsets ($\gamma = 0.85$)	2.323	2.004	2.510	0.987	0.816
Estimated Offsets ($\gamma = 0.7$)	2.341	1.976	2.551	1.061	0.844

Table 4.1: Phase Offset Estimation, Ramp-like Profile

error on the height profile, defined as:

$$\varepsilon_{\mathbf{h}} = \frac{\|\hat{\mathbf{h}} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2} \quad (4.14)$$

using the true offsets is equal to $\varepsilon_{\mathbf{h}} = 9.3259 \times 10^{-4}$, while the normalized reconstruction square error using the estimated offsets is equal to $\varepsilon_{\mathbf{h}} = 1.0105 \times 10^{-3}$. As a second experiment, the same height profile and same systems parameters are considered, but in a more noisy situation, reducing $\gamma = 0.7$. The results of the phase offset estimation are shown in the fourth row of table 4.1. The normalized reconstruction error on the height profile, using the estimated offsets is equal to $\varepsilon_{\mathbf{h}} = 2.0 \times 10^{-3}$, while the normalized reconstruction error using the true values

4. PHASE OFFSET ESTIMATION

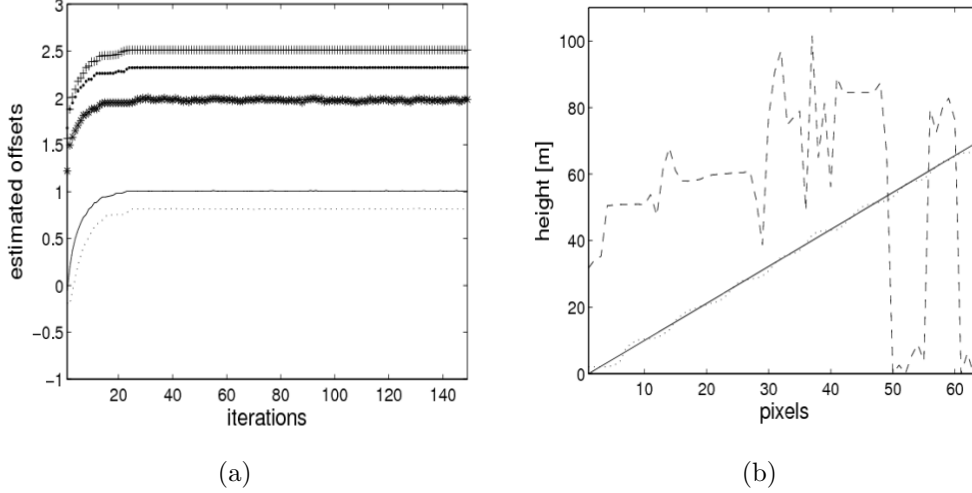


Figure 4.3: (a) Phase offset estimation with $\gamma = 0.85$. Point, star, plus, solid, and dotted lines represent 1st, 2nd, 3rd, 4th, and 5th phase offset estimation; (b) cut of the height profile (solid line), and of the reconstructed height profile using the estimated phase offsets (dotted line and without the phase estimation (dashed line

is equal to $\varepsilon_{\mathbf{h}} = 1.7 \times 10^{-3}$. As expected, the performances for $\gamma = 0.7$ are worse respect to the case with $\gamma = 0.85$.

As a third experiment, a different height profile with the same systems conditions is considered. It consists of a Gaussian profile with a maximum height of 70 m (figure 4.4). The coherence is set to $\gamma = 0.85$. The offset estimation results are summarized in the third row of 4.2 and represented in figure 4.5(a).

Interferogram	1st	2nd	3rd	4th	5th
True Offsets	2.3	2	2.5	1	0.8
Estimated Offsets ($\gamma = 0.85$)	2.324	2.010	2.515	1.006	0.819
Estimated Offsets ($\gamma = 0.7$)	2.389	2.023	2.576	1.101	0.908

Table 4.2: Phase Offset Estimation, Gaussian Profile

Once again, after the estimation of the phase offset, $\hat{\psi}$ are used for the MAP multichannel reconstruction of the profile. The normalized reconstruction error

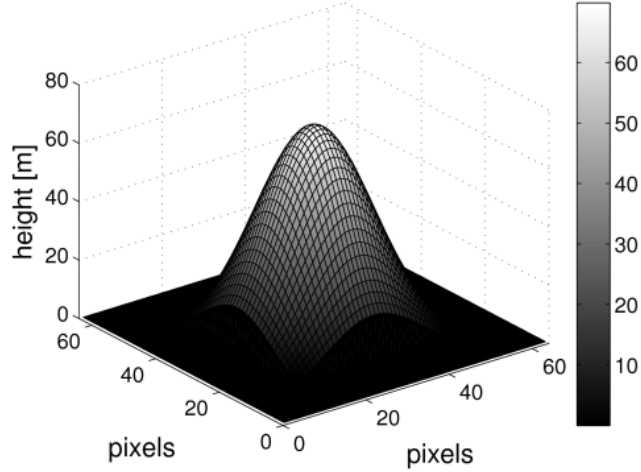


Figure 4.4: Phase offset simulated experiment - Gaussian profile - Reference height profile

on the height profile, using the estimated offsets is equal to $\varepsilon_{\mathbf{h}} = 3.1 \times 10^{-3}$, while using the true values is equal to $\varepsilon_{\mathbf{h}} = 3 \times 10^{-3}$. The MAP reconstructions of the height profile using the true offsets, the estimated offsets and without using the offsets estimation procedure are represented respectively in figures 4.5(b), 4.5(c) and 4.5(d). As a fourth example, the Gaussian height profile of the previous example 4.4 is considered but with coherence $\gamma = 0.7$. The offset estimation results are summarized in forth row of 4.2. Once again, the results for $\gamma = 0.7$ are worse respect to the Gaussian case with $\gamma = 0.85$.

The obtained results, in all the simulations, prove the effectiveness of the method, and assess the overall quality of the height estimation procedure.

4.5 Real Data Experiments

The offset estimation procedure presented in section 4.3 has been tested also on real data. An ENVISAT data set acquired over the area of Las Vegas has been investigated. The acquisition parameters and baselines are the same of the previous simulation examples, already introduced in section 4.4. Each interferogram is composed of 1596×796 pixels, among which $M = 1400$ high coherence pixels

4. PHASE OFFSET ESTIMATION

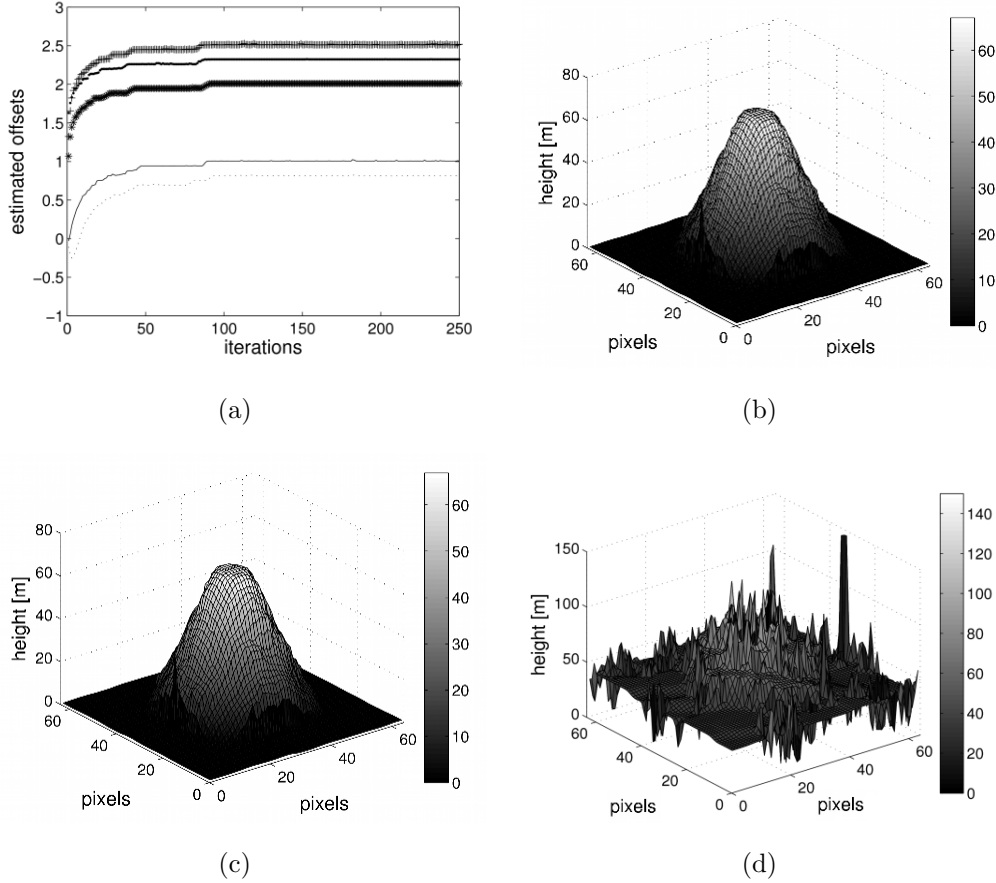


Figure 4.5: (a) Phase offset estimation, with $\gamma = 0.85$: point, star, plus, solid, and dotted lines represent 1st, 2nd, 3rd, 4th, and 5th phase offset estimation; (b) reconstructed height profile, using the true offset; (c) reconstructed height profile, using the estimated offset; (d) reconstructed height profile without using the phase offset estimation procedure

have been selected. In order to select the M pixels needed for the phase offset estimation procedure, a threshold of $\gamma \geq 0.68$ on each interferogram has been set. With this experiment we show how an optimal estimation of the phase offsets can be used to improve the resolution of an available SRTM DEM, which is known to be provided with a low resolution. For this reason an SRTM DEM of the same area is considered.

We first estimate the phase offsets of the 5 interferograms (figure 4.6(a)) using the

above described procedure. The number of iterations needed to the convergence of the phase offset estimation is of about 50 iterations.

Then using the estimated phase offset, we perform the height reconstruction using the MAP multichannel technique. After the height reconstruction, the difference between the obtained height reconstruction of the area under test and the SRTM DEM of the same area is calculated. This difference is shown in figures 4.6. In particular 4.6(b) shows the difference between the MAP reconstruction and the SRTM DEM when the MAP reconstruction is realized ignoring the existence of phase offsets (i.e. without recurring to their estimation). Figure 4.6(c) represents the same difference, when the phase offset estimation is carried out. As it is clear, using the offset estimation technique and the multichannel MAP reconstruction method, we can improve the accuracy of the SRTM DEM. This is more clear if we look at the representation of the residual gradient in figure 4.6(d). In this figure all the added details of our MAP multichannel reconstruction with phase offset estimation respect to the SRTM DEM are highlighted. The presented result is consistent with the one presented in [33], where, working on the same data set, but using a completely different methodology, a comparable improvement of SRTM DEM accuracy is obtained.

4.6 Conclusions

The problem of phase offsets presence in multichannel interferometric SAR data has been discussed in this chapter. The presence of phase offsets on the interferograms is due to several factor, such as processing errors, parallel baseline inaccuracy. The estimation of these phase offsets is mandatory in order to restore the correct relation between the different interferograms used in the MAP (or ML) height reconstruction procedure. A statistical method based on ML technique to estimate these offset has been proposed. In particular the EM algorithm has been used. The estimation procedure has been tested on both simulated and real data. The offset estimation procedure in simulated data case, has been tested with different height profiles and with different noise condition. It has shown good performances in all the examined cases and it has assessed the overall quality of the height estimation procedure. Referring to the real data case, it has

4. PHASE OFFSET ESTIMATION

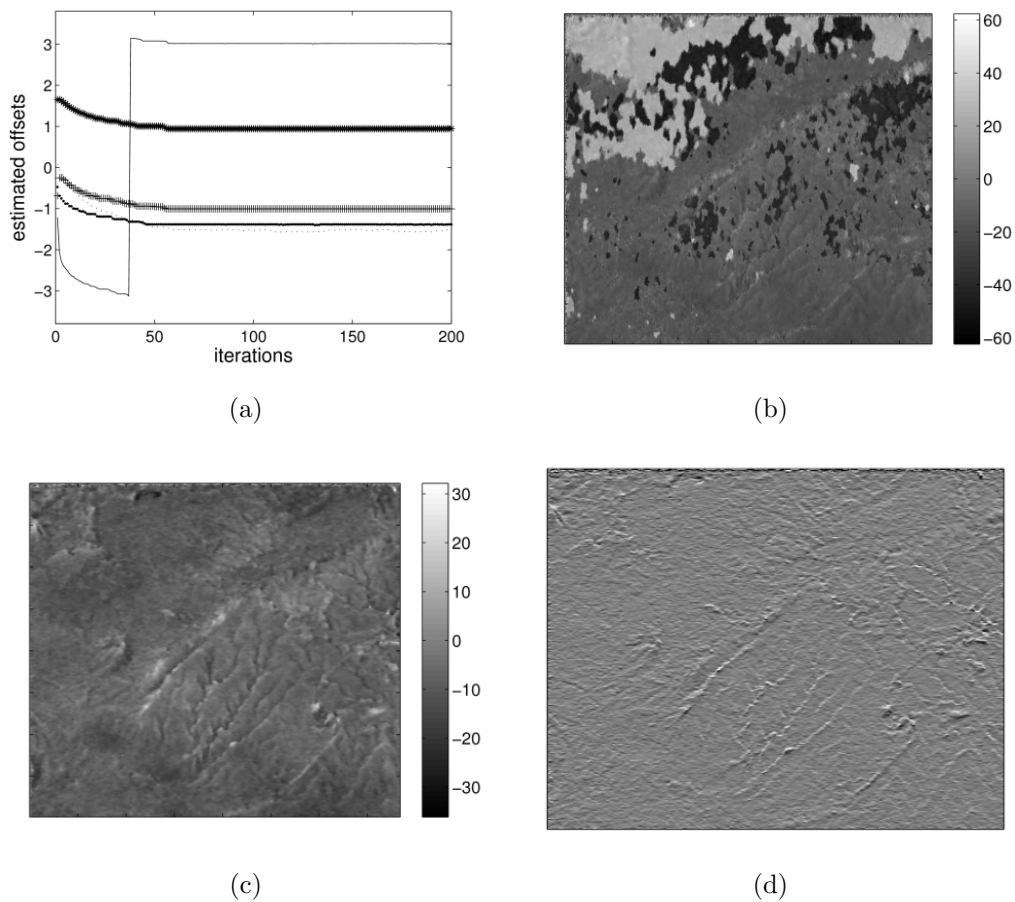


Figure 4.6: (a) Phase offset estimation; (b) difference between our MAP reconstruction and the SRTM DEM, without offsets estimation; (c) difference between our MAP reconstruction and the SRTM DEM, after using phase offsets estimation; (d) residual gradient

been shown that, once the phase offsets have been estimated using the proposed procedure, using the MAP-based height reconstruction algorithm it is possible to successfully improve the accuracy of an existing DEM (SRTM).

5

Multichannel InSAR in Urban Areas

5.1 Introduction

In this chapter the Multichannel InSAR systems in urban areas are discussed. Urban scenarios are difficult to be processed due to the presence of SAR geometrical distortions, layover above all. Two fields of research have been investigated. The first one regards the use of Multichannel InSAR systems for buildings edge detection. An approach based on LGMRF is proposed. The algorithm finds the edges of the image starting from the estimation of the hyperparameters of the LGMRF model. The hyperparameters are seen as an index of the spatial correlation between adjacent pixels. The second research field regards the introduction of auxiliary data, optical data or SAR shadowing, in the MAP multichannel height reconstruction in order to improve the reconstruction in urban areas affected by the layover effect. In the first part of the chapter the layover phenomenology in SAR images of urban areas is described. Then the algorithm to detect edges in urban areas is described. Finally the procedure to improve the height reconstruction in layovered areas is explained.

5.2 InSAR Layover Phenomenology in Urban Scenario

The layover is a geometrical SAR distortion effect typical of very steep height profiles (see section 3.2.1) as the ones present in urban areas. Due to the SAR geometry (the SAR image is available in the slant range coordinate), in urban areas the echo of roof anticipates the echo coming from the ground close to the vertical wall of the building, since the slant range of the roof-wall edge is smaller than the other one. In urban scenarios, layover always affects buildings: echoes from flat earth near building, vertical wall and roof are always overlapped. The geometry of a building subject to layover is shown in figure 5.1. It can be

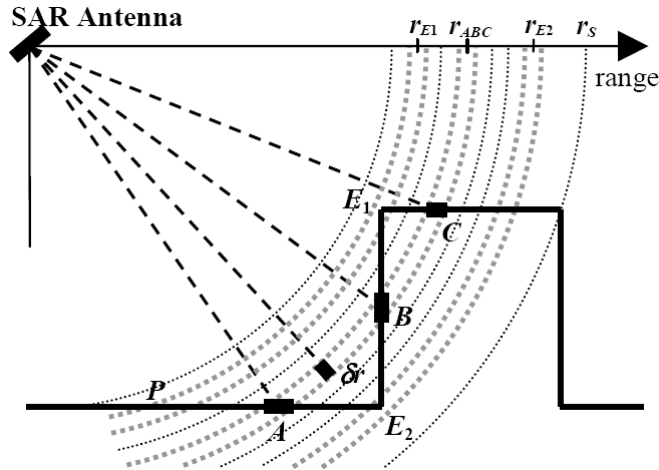


Figure 5.1: Layover Building Geometry - Representation of a building affected by the layover

observed that the roof-wall edge E_1 is seen from the SAR sensor at a range r_{E1} smaller than the wall-ground edge E_2 at range r_{E2} . All the range between r_{E1} and r_{E2} is characterized by multiple contributions in the same range resolution cell, whose spatial resolution is δr . In particular, in the case reported in 5.1, the three small areas A, B and C give contribution in the same range cell positioned at a distance r_{ABC} from the sensor. Moreover, it has to be considered that areas positioned in the dihedral formed by ground and vertical wall provide single scattering contribution (direct single path reflection toward the SAR antenna) and

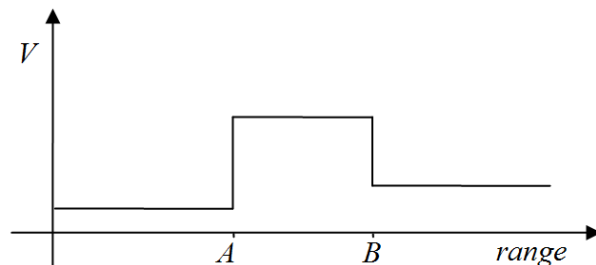
5.2 InSAR Layover Phenomenology in Urban Scenario

multiple scattering contribution. This latter can be summarized as a dihedral corner reflector whose contribution to the SAR final image is localized at a range r_{E2} .

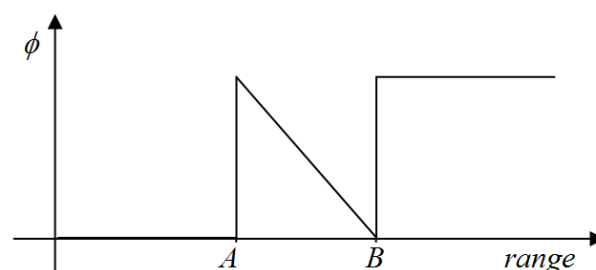
The SAR received signal is characterized by a single contribution from the ground (at the left of point P in figure 5.1) at slant ranges smaller than r_{E1} , by the three contributions given by ground (A), vertical wall (B), and roof (C) in the layover region at slant range between r_{E1} and r_{E2} , by a corner reflector contribution positioned at the range r_{E2} , and then by a single contribution given by the roof at slant ranges larger than r_{E2} . After, there is the shadowing effect generated by the building at ranges larger than r_s . In this case, the signal corresponding to the layover region depends both on amplitude (related to the material, roughness, viewing angle) and phase (related to the object and to the distance between sensor and object) of each single contributions.

Vertical wall contribution is normally stronger than the other two contributions (note that this is not the only possible case, but is what we focused on), basically for the presence of window shutters and balconies: their effect is similar to the one due to small corner reflectors. Moreover, often there are metallic parts that reinforce the backscattered signal.

In this hypothesis the layover phenomenon in SAR images produces two particular effects. SAR intensity image u , in layover area, is characterized by high reflectivity. SAR interferogram ϕ is characterized by a ramp with two discontinuities, one at the beginning of the layover, one at the end of the layover area, with a ramp-like trend between them. This can be better understood, considering figure 5.2. Consider first figure 5.2(a). Suppose it represents the intensity $V = |u|^2$ of the SAR image. Before point A there is a certain reflectivity due to the ground floor. The area between point A and point B represents the layover area, with an higher reflectivity. Finally after point B the roof of the building, with a third reflectivity is represented. The layover effect in the interferogram can be seen in figure 5.2(b): before point A the phase ϕ is constant and represents the ground floor height. The area between point A and point B represents the layover area (a ramp with two discontinuities), after point B phase is constant again and represents the roof height.



(a)



(b)

Figure 5.2: (a) Intensity image of a building affected by layover, (b) Interferogram of a building affected by layover

5.3 Urban Edge Detection

In this section edge detection in urban areas is investigated. Edge detection is an important tool in many image processing spheres, such as segmentation, feature extraction and classification. An edge can be defined as a change in the reflectivity or the intensity of the image pixels [43].

Most of the edge detection techniques are based on the evaluation of the gradient (in practice they compute the different pixels intensities). The Sobel edge filter is one of the most effective and used edge detector based on gradient techniques [44], [45]. These techniques, however, need a previous smoothing and/or filtering, to work correctly since differential operations are sensitive to noise.

In case of Synthetic Aperture Radar images, the edge detection based on gradient techniques do not exhibit good results, essentially due to the multiplicative nature of SAR noise, i.e. the speckle [37]. Due to the multiplicative characteristic of the speckle, gradient based techniques detect more false edges in high reflectivity areas than in low ones. This lead to the problem of a non Constant False Alarm Rates (CFAR). Improved detection can be obtained by speckle filtering [16] the image prior to edge detection, with some trade-off in resolution and computational complexity.

Attempts to deal with the problem of speckle in SAR edge detection has led to the development of statistical or ratio-based techniques for edge detection, that guaranty CFAR [46] [47]. These techniques have shown to be optimal only in the case of mono-edge case (i.e. a single edge can be found within the analysis window, which may be quite large to suppress the speckle). An improvement of these approaches, that makes possible the multi-edge detection, is obtained using windows of different size [48]. Stochastic methods based on Markov Random Field and Simulated Annealing have shown to be effective and to provide good results [49].

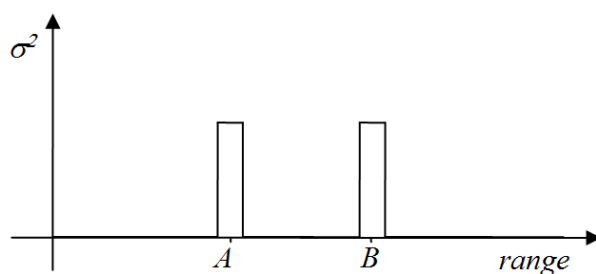
In a urban scenario we are interested in edges detection in order to discriminate buildings form the surrounding areas. Often, the edges corresponding to the transition from different parts of the buildings (vertical wall, ground floor and roof) can be discriminated only if they show very different intensity. Otherwise the edge detection can easily fail using traditional techniques. An interesting work on edges detection in urban ares is [50], where the buildings edges are detected using the fusion of SAR data and Optical data.

The approach we proposed is a stochastic approach based on the idea of finding edges in a urban scenario image by modeling can be seen as an index of the spatial correlation of the pixels. The estimation of the edge is carried out by estimating the hyperparameters.

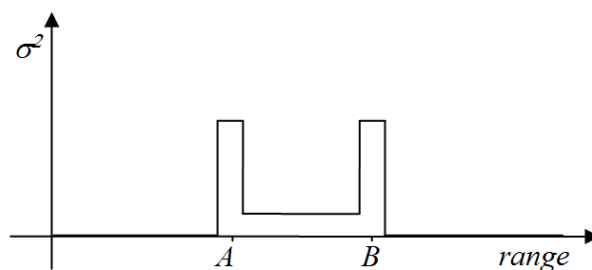
The edge estimation is performed starting from interferometric phase signals, as finding the edges in urban areas which corresponds to the localization of different reflectivity areas in the intensity image V (ground floors, vertical walls, roofs), is equivalent to find the edges of InSAR images (due to phase jumps between adjacent pixels). Consider the building affected by layover of figure 5.2. The edge

5. MC INSAR IN URBAN AREAS

detection on the two images finds in both cases two edges: one at the beginning of the layover area in point A (where the vertical wall starts), one at the end of the layover area in point B (where the vertical wall ends). The detected edges in case of SAR intensity image V (figure 5.2(a)) and in case of InSAR data ϕ (figure 5.2(b)) are shown respectively in figure 5.3(a) and in figure 5.3(b). Moreover,



(a)



(b)

Figure 5.3: (a) edges detected from the intensity image V of figure 5.2(a), (b) edges detected from the interferogram ϕ of figure 5.2(b)

in InSAR systems exploiting the Multichannel approach it is possible to obtain more independent data, that if correctly combined can help in providing a good solution, without having a loss of resolution.

5.3.1 InSAR Urban Edge Detection

The idea is to perform, starting from interferometric data of a scene, the edge detection based on different height of the pixels instead of on the difference reflectivity of the pixels. Consider $\mathbf{h} = [h(1)h(2) \cdots h(S)]^T$ the height of the pixels image. We have seen in chapter 3 that \mathbf{h} is not known and it has to be estimated starting from the wrapped measured interferometric data. The unknown heights can be modelled by some statistical distribution. An effective way to model the unknown pixels height is to use LGMRF (see section 3.4.2). In the LGMRF model the energy function is characterized by a vector of local hyperparameters $\boldsymbol{\theta}$ (3.33).

On the vector of hyperparameter $\boldsymbol{\theta}$, the proposed edge detection is based. As said in section 3.4.2, the hyperparameter $\theta_{s,j}$, where s and j are two neighbor pixels, can be seen as an index of the spatial correlation of neighboring pixels. High value of $\theta_{s,j}$ means that the probability that two pixels $h(s)$ and $h(j)$ assume very different height values is high. Low value of $\theta_{s,j}$ means that the probability that $h(s)$ and $h(j)$ are very different is small. For the edge detection problem, this practically means that an high value of $\theta_{s,j}$ corresponds to a transition between different height values (an edge between s and j); a low value of $\theta_{s,j}$ means no height transition (no edge). It is important to note that $\theta_{s,j}$ is not known but it has to be estimated. The estimation can be performed exactly in the same way as explained in section 3.4.2. Estimation of $\theta_{s,j}$ corresponds to perform an edge detection.

Once the hyperparameters have been estimated, a comparison with a threshold is made. Since an hyperparameter bigger then 1 means that the two neighbor pixel can have a different height value, the threshold is set to be 1. If $\hat{\theta}_{s,j}^2$ is bigger than the threshold than an edge is detected between pixel s and pixel j . Otherwise no edge is detected.

In order to improve the performances of the proposed algorithm, multichannel InSAR systems can be exploited (section 3.4). The use of a multichannel InSAR system has a two folded effect: first, using more independent data, a clearer result (less noisy, i.e. less false alarm) is reached without loosing in resolution. Secondly, the use of multichannel data avoids problems related to the wrapping

5. MC INSAR IN URBAN AREAS

of the interferometric phase. Suppose the height of a building is such that, according to relation (3.15), it corresponds to a phase value multiple of 2π . This value will wrap and it will provide a phase value equal to 0. In this case, using a single interferogram the edge detection would fail, providing no edges. Using a second interferogram, obtained using a different channel, the same height value will be mapped into a different phase value (no more multiple of 2π). Combining the two interferograms, the algorithm will perform correctly the edge detection. Using this approach the a posteriori distribution used for generating samples in the hyperparameter estimation step is given by (3.35).

5.3.1.1 Numerical Experiments on Simulated and Real Data

As a first experiment we consider a simulated profile (64×64 pixels), consisting of 5 buildings of different height (see figure 5.4(a)). We suppose that there is no layover and only the roof can be seen (no vertical wall). An interferogram is simulated using SRTM parameters, with a coherence $\gamma = 0.8$ and frequency $f = 5$ GHz (figure 5.4(b)). First, we apply a gradient based edge detection on the interferogram. The results are shown in figure (5.4(c)). Due to the noise characterizing the interferogram, the gradient based edge detector fails to correctly locate the edges. Then we apply our edge detection algorithm to the interferograms. the results are shown in figure 5.4(d). As it can be seen the edge detection correctly finds the edges.

The obtained results can be improved using the Multichannel edge detection. We simulated for this purpose a second interferogram with a coherence $\gamma = 0.8$ (figure 5.4(e)) and frequency $f = 9$ GHz. As expected the performances improved. Note that the improvements have been obtained without filtering or reducing the spatial resolution. As expected the performances improved. Note that the improvements have been obtained without filtering or reducing the spatial resolution.

As a second experiment we considered a simulated profile (64×64 pixels), consisting in a building characterized by layover phenomenon. Figure 5.5(a) shows the simulated SAR intensity image without noise. There is an high reflectivity area corresponding to the layover area (brighter area), two areas of similar intensity (roof and ground floor), and a dark area representing the shadowing. Note

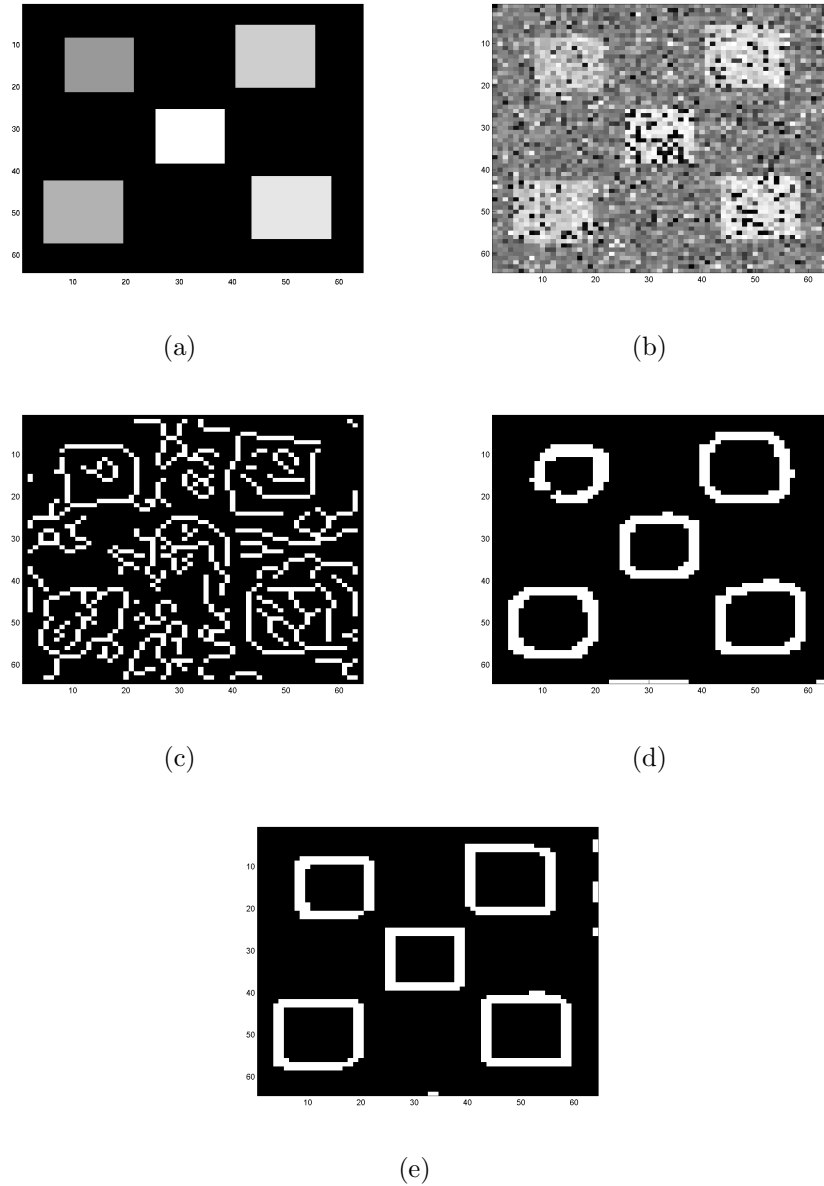


Figure 5.4: (a) Considered Height Profile, (b) simulated interferogram, (c) interferogram edge detection using gradient based techniques, (d) singlechannel edge detection using the proposed algorithm, (e) multichannel edge detection using the proposed algorithm

5. MC INSAR IN URBAN AREAS

that the roof is the area between the beginning of the layover and the shadowing. We generated multiplicative noise (speckle) in order to simulate noisy SAR intensity data (fig 5.5(b)). Then we generated the noisy interferograms of the scene (figures 5.5(c) and 5.5(d)). To generate the interferograms we used ENVISAT parameters with two baselines $[230m \quad 430m]$, with a coherence $\gamma = 0.6$ and frequency $f = 5\text{GHz}$. We generated two azimuth looks for each baseline, allowing a total number of 4 channels. To generate the interferometric phase we considered a building having an height of 30m. This height value is such that the wrapped phase, in the case of the second baseline, corresponds to 2π , falling in the case discussed before (figure 5.5(d)). In this conditions the singlechannel edge detection would fail.

We performed the multichannel edge detection on the interferometric data, obtaining the results of figure 5.5(f). It can be seen how the algorithm is able to well locate the expected edges: the one due to the vertical wall, and the one due to the roof. We performed the edge detection on the combination of the 6 available SAR intensity image, using a gradient based technique. As expected the detection recovers many false alarm in high reflectivity areas than in low one (figure 5.5(e)).

Finally, for the real data experiments, we considered a real data set of the city of Dresden. We used 2 L-Band E-SAR interferograms obtained with two baselines $[10m \text{ and } 40m]$. We applied the multichannel edge detection algorithm on this data set. Figure 5.6(a) shows the SAR intensity image of the considered scene. Figure 5.6(b) shows one of the two available interferograms (smallest baseline). The multichannel edge detection has been performed on the interferograms, providing the edges of Figure 5.6(d). Figure 5.6(c) shows a gradient based edge detection on the three available intensity images. Many false edges have been detected in this case. The good performances of the proposed method can be easily seen, especially if the result are compared with an optical image of the zone provided by Google Earth (figure 5.6(e)). The edges found using the proposed approach agree with the real buildings edges. Buildings edges are very well recognized and there are not many false alarm.

5.3 Urban Edge Detection

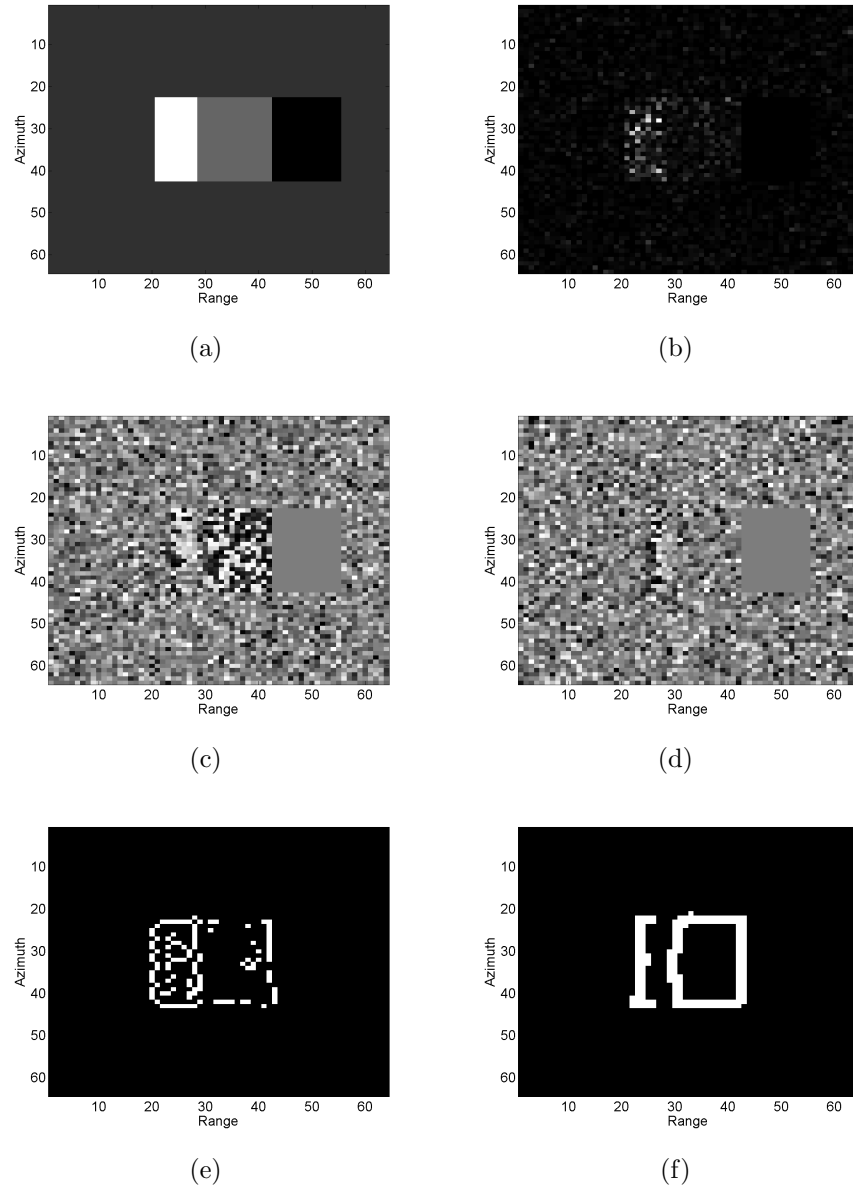


Figure 5.5: (a) Considered Profile, (b) Noise free intensity Image, (c) Simulated interferogram (baseline 230m), (d) Simulated interferogram (baseline 430m), (e) Gradient based Edge detection of the 6 available intensity images, (f) multichannel edge detection using the proposed algorithm

5. MC INSAR IN URBAN AREAS

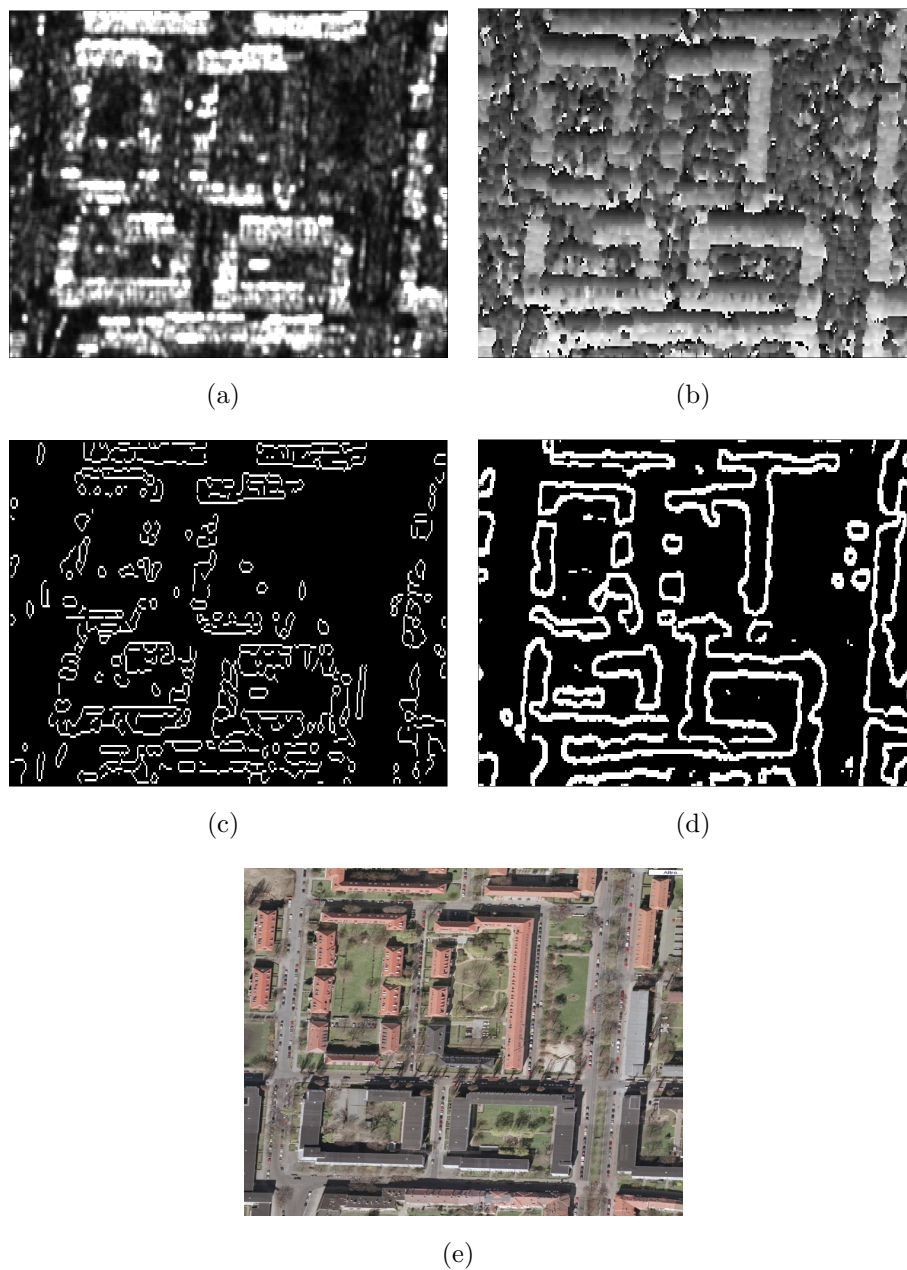


Figure 5.6: (a) One of the three available intensity images, (b) One of the two available interferograms (baseline 10m), (c) gradient based edge detection using the three available intensity images, (d) multichannel edge detection using the proposed algorithm, (e) Optical image of the area (thanks to Google)

5.4 MAP multichannel DEM reconstruction using Auxiliary Data

In this section the introduction of auxiliary data in the MAP multichannel height reconstruction (section 3.4.2) in order to improve the reconstruction in urban areas is investigated.

As previously said, urban areas are a difficult to be processed due to the presence of geometrical distortions, especially for the effect of the layover (section 5.2). In case of layovered areas, the classical phase unwrapping methods fail to retrieve the correct solution. In fact the two discontinuities, one at the beginning and one at the end of the layover (figure 5.2), are often bigger than π . Furthermore the interferometric phase is corrupted by interferometric noise (characterized by low coherence areas). These conditions make the phase unwrapping using a single interferogram an impossible task [51].

A dual-channel InSAR system, using the MAP approach (see section 3.4.2) can overcome this problem, providing a good reconstruction of the area, both of the two discontinuities and the ramp between them, also in the case of a narrow layover area (in case of wide view angles or in case of vertical wall not orthogonal to the floor). Even if the dualchannel MAP results are good, the idea is to improve the height reconstruction in particular on the discontinuities that layover implies. To do that, some auxiliary information are introduced in the multichannel MAP height estimation procedure. How to take advantage of these information is explained in the next section.

A key part of the implementation of a successful MAP multichannel approach is a good knowledge of the local hyperparameters. Auxiliary data can be used to help the generation of hyperparameters map localizing discontinuities in the scene. In order to improve height reconstruction estimation, in the case of a vertical wall orthogonal to the ground (typical of urban scenario), we considered two different types of auxiliary external data: the information carried by an optical sensor and the shadowing areas of SAR images.

Optical images are usually characterized by multiple bands, and even using a single date acquisition they provide significant classification performance. Unfortunately, the availability of optical images of a given area strongly depends on

5. MC INSAR IN URBAN AREAS

weather conditions, while SAR images are not largely affected by atmospheric phenomena.

The geometrical model for optical image acquisition is completely different compared to the SAR one and is based on the optical center. Each point of the image is obtained by the intersection of the image plane and the line joining the 3-D point and the optical center. The length of a surface in optic image is related to viewing angle it covers (α in figure 5.7).

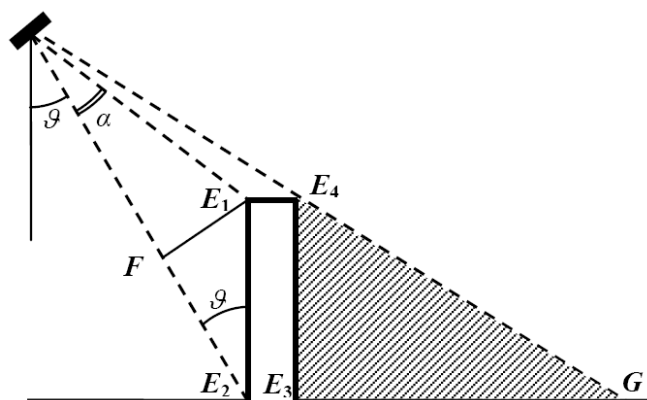


Figure 5.7: Optical and SAR geometry - Representation of a building seen from a SAR and from an optical sensor

Anyway, even if optic images permit an easier object and boundaries identification, they are not accurate enough to allow digital elevation model generation (basically due to the sensor incoherence and to uncontrolled build geometry). Fusion of SAR and optical data is possible only in case of a good co-registration of the signals [52] (i.e. images have to be superposable), which is an easy task to be performed if SAR and optical sensors share the same point of view and look at the same region, otherwise a transformation has to be applied.

The basic idea of registration is to use ground control points, which we expect to have the same position in both images. In SAR images suitable points are those at the bottom of vertical wall of each building (point E_2 of figure 5.7), where scene configuration traces a corner reflector. The corner reflector is characterized by a strong return and it is easy to be detected. The building vertical wall signal finishes in this point in the SAR case, while it starts in this point in the optical

5.4 MAP multichannel DEM reconstruction using Auxiliary Data

domain. Although in SAR domain (i.e. slant range domain) we don't know the vertical wall signal length (which is the layover length, FE_2 in figure 5.7), in optical domain it is easy to measure it, once identified. To identify it in optical domain optical segmentation techniques can be used [53].

The idea is to compute this length in optical domain (FE_1 in figure 5.7) and therefore estimate layover length (FE_2 in figure 5.7) using the trigonometry identity

$$FE_2 = \frac{FE_1}{\tan \theta} \quad (5.1)$$

Note that $FE_2 \approx r_{E_2} - r_{E_1}$ of figure 5.1.

In this way we are able to identify and locate discontinuities in SAR interferogram and therefore to improve the hyperparameters map which is then used for multichannel MAP reconstruction.

The other auxiliary information sources that can be exploited are the shadowing areas (see section 3.2.1). Shadowing areas are always present in urban scenarios; they are complementary to layovers areas and occur when a structure makes certain ground areas not covered by radar beam (E_4G in figure 5.7); these zones produce no echoes and appear dark (zero amplitude signal) in the slant range domain, with any changes due solely to system noise, sidelobes, and other effects normally of small importance. As stated for the optical case, shadowing area length can be related to the height of the building and to viewing angle via the trigonometric identity

$$E_3E_4 = E_4G \cdot \cos \theta \quad (5.2)$$

allowing us to estimate layover area length by:

$$FE_2 = E_1E_2 \cdot \cos \theta = E_3E_4 \cdot \cos \theta = E_4G \cdot \cos^2 \theta \quad (5.3)$$

to properly build an hyperparameters map. Clearly this is valid in case of building with the same height on the two sides. To localize the layover area on the interferogram we use SAR amplitude image, where the location of point E_2 (the end of the layover area) in figure 5.1 is characterized by an easy to find strong echo due to corner reflector shape.

Note that of an improvement of the hyperparameters map also the edge detection technique of section 5.3 would be positively affected. So the proposed method

5. MC INSAR IN URBAN AREAS

can be applied also for the edge detection problem.

5.4.1 Simulated Experiments

In order to assess the validity of the procedure presented before, in this section some results on simulated data are presented. A building of 30m height with a flat roof is considered. Two interferometric phase signals are generated using 2 baselines (230m e 430m), coherence $\gamma = 0.8$, range distance $R = 700$ km, wavelength $\lambda = 0.0566$ m and look angle $\theta = 40$. Two azimuth looks are generated for each baseline, so allowing $N = 4$ independent channels. In figure 5.8 the reference pro-

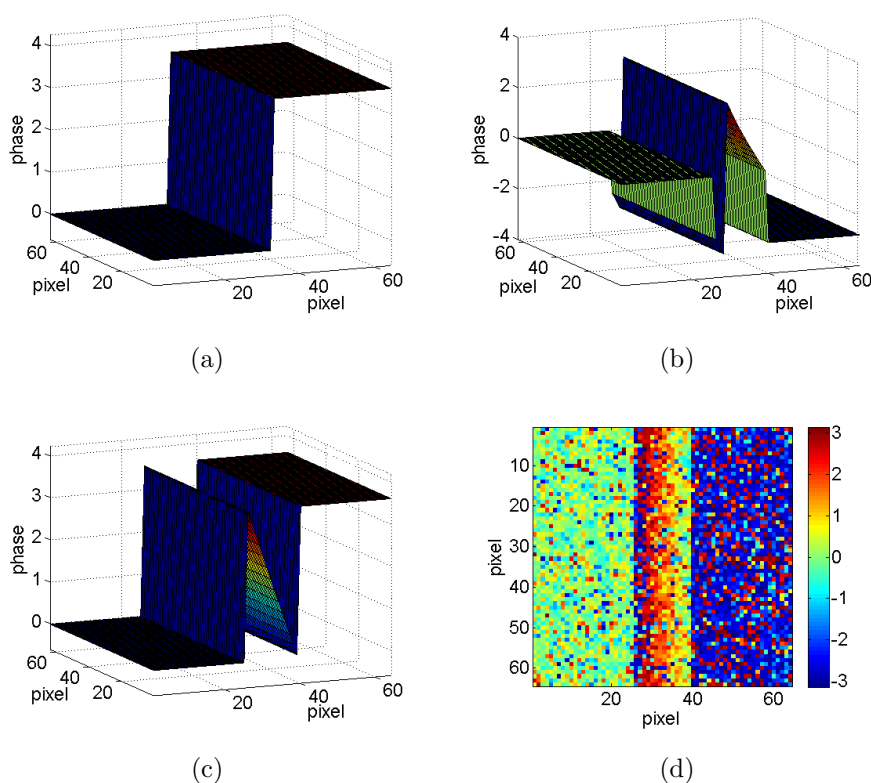


Figure 5.8: (a) Reference profile (230m baseline), (b) Interferometric ramp, (c) Unwrapped reference ramp, (d) Interferometric ramp affected by noise

file, the interferometric reference ramp, the interferometric reference unwrapped

5.4 MAP multichannel DEM reconstruction using Auxiliary Data

ramp and one of the four noisy interferograms, all in case of 230 m baseline are shown. It is important to notice that even the first interferogram (smaller baseline) is ambiguous. In fact the height discontinuity of 30m corresponds respectively to 3.6rad and 6.3rad. In these conditions (discontinuity bigger than π) a classical phase unwrapping method based on a singlechannel would fail. The dual channel MAP reconstruction is applied to the considered data set. We use the hyperparameter vector estimated using the EM algorithm, as in section 3.4.2. The result is shown in figure 5.9(a). The normalized reconstruction square error is $\varepsilon_{\mathbf{h}} = 3 \times 10^{-3}$. Although the reconstruction, considering the limited number of available data ($N = 4$ channels), is good, the idea is to improve its quality, especially on the discontinuities.

The procedure based on the introduction of auxiliary external data is then applied. In particular, the auxiliary optical data are exploited. To do this, data provided by a an optical sensor located in the same position of the SAR are simulated, with the same look-angle θ . After the co-registration of the data (optical and InSAR), the optical information is introduced in order to correct the a priori term and perform the MAP estimation. Note that the same can be done starting from the shadowing present in the SAR images. Figure 5.9(b) shows the MAP reconstruction of the building obtained using auxiliary external data. Compared to the reconstruction obtained without the additional data, the reconstruction appears less noisy and more precise especially on the discontinuities. The normalized reconstructed error in this case decreases to $\varepsilon_{\mathbf{h}} = 9 \times 10^{-4}$. The improvements can be better appreciate if we compare the geocoded reconstructions (figures 5.9(c) and 5.9(d)). Figure 5.9(e) shows a cut of the reconstructed and georeferenced profile points (triangles) obtained exploiting auxiliary data, overlapped to the reference profile (solid line). The georeferenced profile is obtained by transforming the reconstructed profile from the coordinate system slant range - azimuth - height to the coordinate system ground range - azimuth - height. It can be noticed how accurate the vertical wall estimation is.

It can be also noted the absence of reconstructed points in the ground and roof regions involved in layover phenomenon; this is due to the hypothesis of a bigger vertical wall contribute which leads us to reconstruct only this part of the layover area. To avoid this, and to obtain a reconstruction of the roof and of the ground

5. MC INSAR IN URBAN AREAS

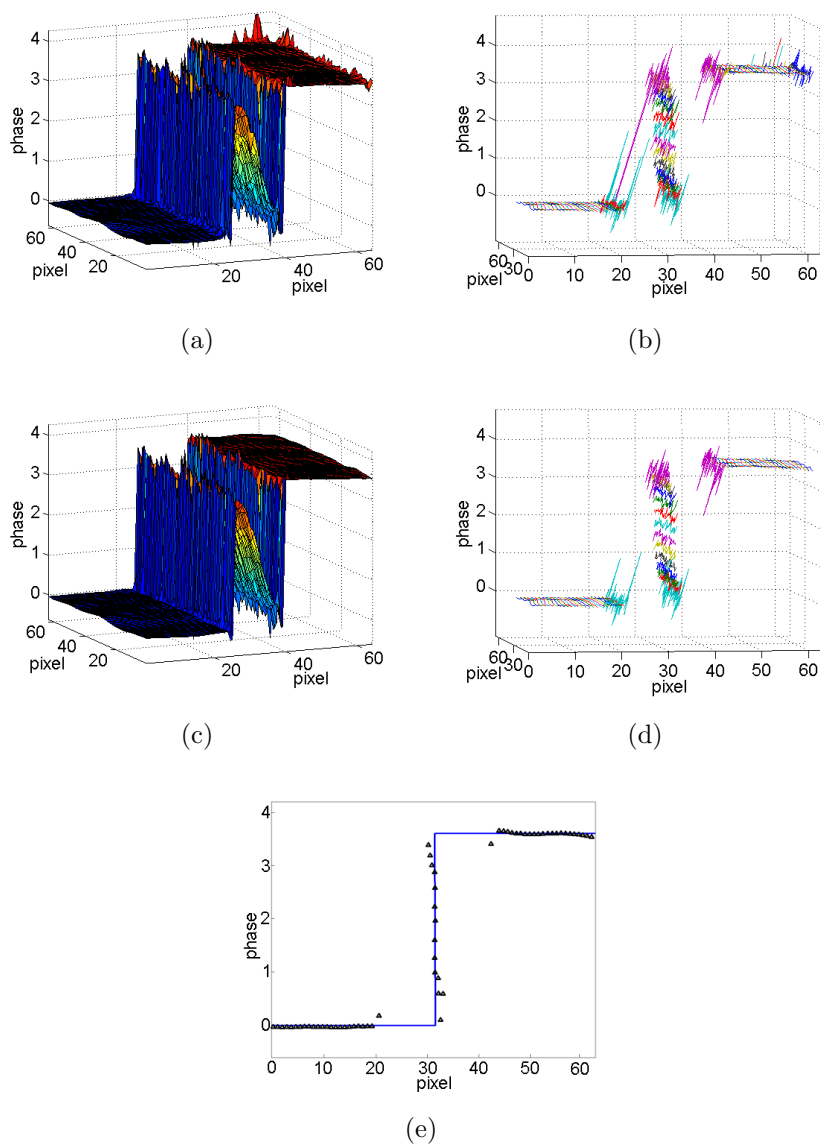


Figure 5.9: (a) MAP 3D reconstruction, (b) MAP 3D reconstruction with auxiliary information, (c) georeferenced MAP 3D reconstruction, (d) georeferenced MAP 3D reconstruction with auxiliary information, (e) cut of georeferenced MAP 3D reconstruction with auxiliary information profile (triangles) overlapped to reference one (solid line)

regions involved in the layover, the SAR tomography approach can be used. A preliminary study on this technique has been conducted (see chapter 7). Some final words about robustness on imperfect estimation of auxiliary data information: first note that the estimation procedure is carried out using MAP approach, and auxiliary data are used to improve the obtained estimation, in particular on discontinuities. In this context, the proposed algorithm is able to take into account small (few pixels) misregistration errors (and/or segmentation errors using optical data). Otherwise, in case of larger errors, the algorithm does not exploit the auxiliary data providing a final solution that agrees with the MAP reconstruction. Note that only one optical image is used, so that the length obtained by FE_1 is only a rough indication of build height which can be validated only after fusion with SAR interferometric data.

5.5 Conclusions

Multichannel InSAR systems in urban areas have been discussed in this chapter. In the first part a new approach based on multichannel InSAR systems for the edge detection in urban areas has been proposed. The algorithm is a stochastic approach based on Markov Random Field, in particular based on Local Gaussian MRF (LGMRF). The approach is based on the estimation of the local hyperparameters of the LGMRF model, which can be seen as an index of spatial correlation between adjacent pixels. The proposed algorithm overcomes the problems typical of SAR edge detection, due to the multiplicative behaviour of noise. We have tested this approach on simulated data and we obtained good results compared to classical gradient based edge detectors. The algorithm has been also tested on a real data set showing the effectiveness of the method. In the second part a methodology to improve the MAP multichannel height reconstruction in urban areas affected by layover has been presented; the approach is based on the combination of the MAP multichannel approach with auxiliary data. In particular the presented method has been developed to correctly manage high phase discontinuities that layover implies. Additional information have been obtained by optical data or from SAR shadowing areas. The proposed method

5. MC INSAR IN URBAN AREAS

has been applied on simulated data. The results obtained on simulated data have shown the effectiveness of the method.

6

Fast InSAR Multichannel Phase Unwrapping

6.1 Introduction

In this chapter a new approach for the multichannel phase unwrapping problem is presented. Statistical approaches, especially the Bayesian ones, have proved to be effective dealing with noisy data and big discontinuities. However, the Gaussian Markov Random Field with local parameters (section 3.4.2) can be time consuming and computationally heavy due to the a priori model and to the optimization step. The idea is to overcome these limits by introducing a new a priori model based on Total Variation and optimization algorithms based on graph-cut technique. A set of experimental results on both simulated and real data illustrates the effectiveness of our approach. In the first part of the chapter some recalls on the multichannel phase unwrapping are made. Then the new fast multichannel phase unwrapping approach is presented. Finally, some experiments on simulated and real data are performed showing the effectiveness of the proposed method.

6.2 Fast MAP Multichannel Phase Unwrapping

The LGMRF model used in [39] (see section 3.4.2) is a model well adapted to any kind of profile (smooth, non-smooth and big discontinuities) thanks to its local behavior. It has been proved that this model combined with the ICM optimization step gives good results for the height reconstruction problem. However, this approach can suffer of some limits.

First, it needs of estimating a parameter $\theta_{s,j}$ for each couple of pixels s, j which is computational heavy and excessively time demanding. In fact if we consider an image of S pixels, for each pixel s the EM (2.5.1) procedure has to be carried out, until convergence, in order to estimate the hyperparameter θ_s . Secondly, once the hyperparameters $\theta_{s,j}$ have been estimated, to find the MAP solution (3.34) [39] proposes two options: the first option is to use the Simulated Annealing; the second option is to use the ICM providing a *semi-deterministic* minimization approach. In the first case the algorithm would be excessively time demanding, due to the characteristic slowly convergence of Simulated Annealing. In the second case, the semideterministic approach, when very noisy and wrapped interferograms are considered, does not guarantee to reach the global optimum, since the ICM is known to be sensitive to local minima. The previous situation can happen when due to not enough available data, the energy function presents local minima very near to the global one.

The idea exposed in this chapter is to provide a fast and efficient (in term of global optimization) algorithm to unwrap the interferometric phase in the multichannel configuration. To reduce the computational time needed to unwrap the multichannel interferometric phase, we worked in two directions: first, we considered a non-local a priori energy function and secondly, we used optimization algorithms based on graph cuts [54].

In order to use a notation more similar to the notation of graph cuts users (and to computer vision problems section 2.5) we can address the MAP multichannel solution as an Energy Minimization problem. In fact the (3.34) can be written as the sum of two energy term, one related to the data, one related to the pixel

interaction (or constraints):

$$\begin{aligned}\hat{\mathbf{h}} &= \arg \max_h F(\Phi|\mathbf{h})P(\mathbf{h}, \boldsymbol{\theta}) \\ &= \arg \min_h (U(\Phi|\mathbf{h}) + U(\mathbf{h}, \boldsymbol{\theta}))\end{aligned}\quad (6.1)$$

The term related to the data (*data energy*) is:

$$U(\Phi|\mathbf{h}) = -\sum_{s=1}^S \log F(\Phi(s)|h(s)) \quad (6.2)$$

while the term related to the pixel interaction (a priori energy) is, in case of LGMRF:

$$U(\mathbf{h}, \boldsymbol{\theta}) = \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} \frac{(h(s) - h(j))^2}{2\theta_{s,j}^2} \quad (6.3)$$

For our purposes we will work on the a priori energy (6.3), providing a different model for it, and on the optimization step, providing a different way to minimize (6.1), giving arise to the *GC-TV* multichannel approach. It is important to mention the works of Bioucas et al. [55] who proposed graph-cut based optimization algorithms to solve the PhU problem and the work of Costantini [28] based on network programming optimization technique. It has to be underlined that these two approaches belong to minimum L^p norm unwrapping methods. So, they do not exploit statistical properties of the noise present on the data and therefore they are not optimal from the information theoretic point of view.

6.3 Local Energy function

As said in the previous section, the idea is to develop a multichannel phase unwrapping algorithm that can overcome the possible limits of the one proposed in the work of [39] (we will call it ICM-LGMRF multichannel approach). In order to achieve this target, a different a priori energy model is proposed: the *Total Variation* (TV) model. The TV model introduced in 1992 [56] is one of the most used prior model in image processing due to its suitability to different contextual information [57; 58]. In SAR applications, TV model is mainly used for image restoration [59]. Our proposal is to apply it to the PhU problem.

6. FAST MCPHU

The a priori energy corresponding to the discretization of TV can be written [60] as follows :

$$U(\mathbf{h}, \theta) = \theta \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} w_{s,j} |h(s) - h(j)| \quad (6.4)$$

where $w_{s,j}$ depends on neighbor connectivity. $w_{s,j} = 1$ for 4 - *connectivity* and $w_{s,j} = \frac{1}{\sqrt{2}}$ for the 4 diagonal neighbors in the case of 8 - *connectivity*. Note that in this expression, θ is a scalar and not a vector of hyperparameters as in the LGMRF model.

TV model is a non local model. The choose of a non local a priori energy is done in order to make the algorithm faster. A non local a priori model avoids the estimation of local hyperparameters $\theta_{s,j}$. Only one parameter θ for the whole image has to be estimated.

Between the existing non local a priori energy model TV has been chosen. Its main advantage, towards other non local a priori energy model model, is that it does not penalize discontinuities in the image while simultaneously not penalizing smooth functions either [56]. TV, in fact, looks for an approximation of the original noisy image which has minimal total variation but without particular bias to discontinuity or smooth solution. As the TV a priori model is well adapted when dealing with strong discontinuities, it can be used in case of InSAR applications and particularly it well fits to urban areas. So, in order to develop a faster algorithm, we considered an a priori energy, which is not as good as the LGMRF one, since it is not local, but between the non local models, it is the one more adapted to our problem.

Given the TV a priori energy model (6.4), the function to be minimized, to recover the multichannel height reconstruction solution, becomes:

$$\begin{aligned} \hat{\mathbf{h}} &= \arg \max_{\mathbf{h}} F(\Phi|\mathbf{h})P(\mathbf{h}, \theta) \\ &= \arg \min_{\mathbf{h}} \left(-\sum_{s=1}^S \log F(\Phi(s)|h(s)) + \theta \sum_{s=1}^S \sum_{j \in \mathcal{N}_s} w_{s,j} |h(s) - h(j)| \right) \quad (6.5) \end{aligned}$$

In the next section the proposed method to minimize this energy is explained.

6.4 Graph-cut Based Optimization

In recent years, energy optimization with graph-cut has become very popular in computer vision [61; 62]. Graph-cut optimization is successful because the exact minimum or an approximate minimum with certain guarantees of quality can be found. Compared to the classical optimization algorithm, Simulated Annealing [5], it provides comparable results with much less computational time and compared to the deterministic algorithm ICM [6], it avoids the risk of being trapped in local minima solution which can be far from the global one.

Two families of graph-cut based optimization algorithms have been developed in recent years. The first one provides the global optimum of the Markovian energy with some constraints on the prior model [60; 63; 64]. The second one provides a local optimum within a good quality and in more general cases of prior energies [61]. To solve the multichannel phase unwrapping problem, which means find the value of \mathbf{h} that minimizes (6.1), we used for the optimization procedure two different graph-cut based algorithms. The first one belongs to the exact optimization algorithm family. In particular, we used the one proposed by Ishikawa [63]. The second is a non-exact optimization algorithm based on α -expansion move [61].

Suppose $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is a directed graph with non negative edge weights, where \mathcal{V} is the set of vertexes and \mathcal{E} the set of edges. This graph has two special vertexes (terminals) called the *source* r and the *sink* t . Note that in the graph cut community the two special vertex source and sink are commonly denoted by s and t . In this thesis the notation r for the source is used in order to avoid confusion with previously defined quantities (s in this thesis is used to denote the pixels or the sites of an image).

A r - t -cut $\mathcal{C} = \{\mathcal{R}, \mathcal{T}\}$ is defined as a partition of the vertexes into two disjoint sets \mathcal{R} and \mathcal{T} such that $r \in \mathcal{R}$ and $t \in \mathcal{T}$. A cost of this cut is defined as the sum of weights of all edges that go from \mathcal{R} to \mathcal{T} . Figure 6.1 shows a simple graph made of 4 vertexes, two vertexes (x, y) plus the sink and the source (t, r) . The edges are the links between the vertexes. An example of a cut is represented with the dashed line.

The minimum r - t -cut problem is to find a cut \mathcal{C} with the smallest cost. This problem is exactly equivalent to its dual problem which consists in computing the

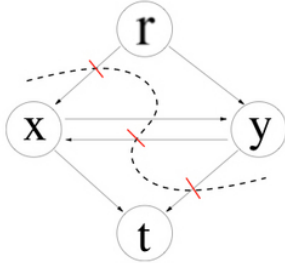


Figure 6.1: Graph representation - Example of a graph construction and of a cut on it (dashed line)

maximum flow from the source to the sink. Many algorithms have been proposed to solve the maximum flow problem with a polynomial time [65]. In our work, we used the maximum flow algorithm given by [66] which is the most adapted to computer vision problems. Let see the two different graph-cut based algorithms that we used for the optimization procedure.

6.4.1 Ishikawa algorithm

The optimization algorithm proposed by Ishikawa [63] belongs to the exact optimization family. Under some hypothesis on the energy to be minimized and if the graph is correctly constructed, Ishikawa algorithm provides the global optimum of the considered energy. Two hypothesis are at the base of Ishikawa algorithm: convexity of the a priori energy and a linear order on the label set. Using the TV model the first hypothesis is satisfied. For the second hypothesis we recall from section 2.2 that the labels are some particular features that have to be assigned to the involved pixels. In the case of multichannel phase unwrapping the labels are the height of the image pixels. We suppose that the labels (the heights) can be represented as integers in the range $\{1, 2, \dots, K\}$, where K is the size of the label sets. This condition satisfies the second hypothesis necessary for the Ishikawa graph construction.

Ishikawa method is based on computing a minimum cut in a particular graph. This graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ contains $S \times K$ nodes (S is the size of the image and K is the size of the label set) denoted by $\{v_{s,k}; s = 1..S; k = 1..K\}$ plus two special

nodes r and t . For each pixel s , we associate K nodes $\{v_{s,k}; k = 1..K\}$ that represent all the possible heights that the pixel s can take. The construction of Ishikawa graph, in case of a 1D image, for legibility, is shown in figure 6.2.

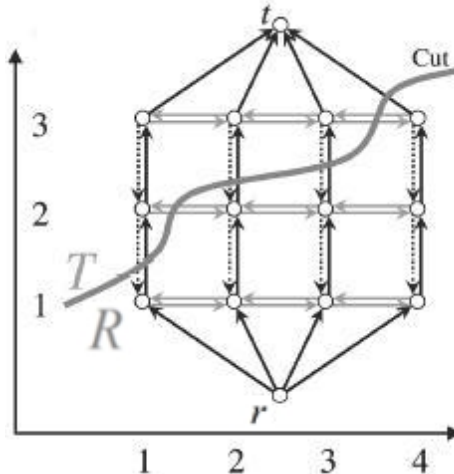


Figure 6.2: Ishikawa Graph Construction - On the axes there are the pixels and the labels; data edges are depicted as black arrows; constraint edges are represented by horizontal arrows and penalty edges are depicted as dotted arrows

Ishikawa graph contains three families of edges $\mathcal{E} = \mathcal{E}_{\mathcal{D}} \cup \mathcal{E}_c \cup \mathcal{E}_{\mathcal{J}}$. $\mathcal{E}_{\mathcal{D}}$ is a set of directed edges called data edges (black arrows of figure 6.2). It represents the data energy term. \mathcal{E}_c is a set of directed edges called penalty edges (dotted arrows of figure 6.2). It ensures that only one data edge is in the minimum cut for each pixel s . Finally, $\mathcal{E}_{\mathcal{J}}$ is a set of constraint edges between all neighbor pixels (horizontal arrows in figure 6.2). It represents the a priori energy term To better understand how to set the edges weight for the multichannel phase unwrapping problem let consider figure 6.3.

The vertex $v_{s,k}$ is the vertex identified by the pixel s and the label k . The cost of the direct edge between vertex $v_{s,k}$ and vertex $v_{s,k+1}$ (black arrow in figure 6.3) is set to be:

$$c(v_{s,k}, v_{s,k+1}) = F(\Phi(s)|k) \quad (6.6)$$

which is the multichannel likelihood value (6.2), calculated for the pixel s and the height value $h(s) = k$. Note that for the two special vertex r and t , the weight of

6. FAST MCPHU

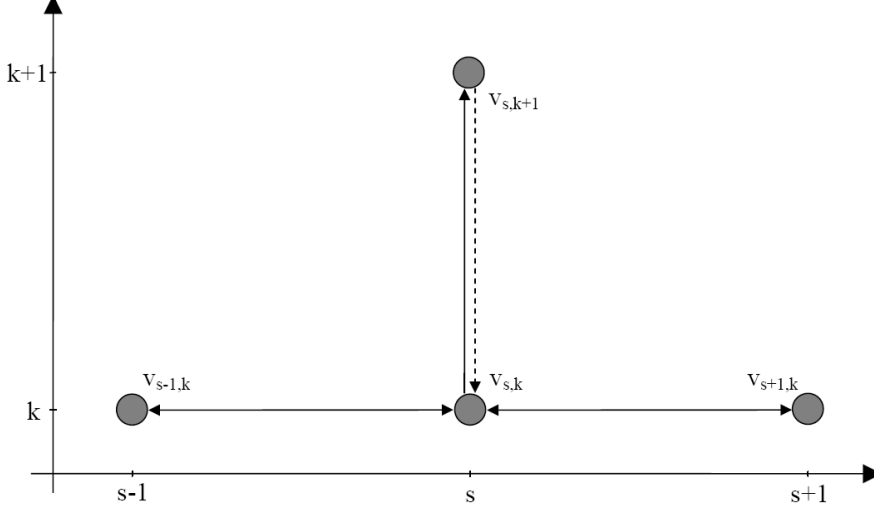


Figure 6.3: Ishikawa edges weights - Representation of the edges for a vertex

the directed edge are, conventionally, set to be:

$$c(r, v_{s,0}) = +\infty \quad (6.7)$$

$$c(v_{s,K}, t) = F(\Phi(s)|K - 1) \quad (6.8)$$

The cost of the penalty edge between vertex $v_{s,k+1}$ and vertex $v_{s,k}$ (dotted arrow in figure 6.3) is set to be:

$$c(v_{s,k+1}, v_{s,k}) = \infty \quad (6.9)$$

Finally the weight of the constraint edges between vertex $v_{s-1,k}$ and vertex $v_{s,k}$ (horizontal arrows in figure 6.3) is set to:

$$c(v_{s-1,k}, v_{s,k}) = 1 \quad (6.10)$$

This last results comes from the hypothesis of TV model (6.4) for the a priori energy. The results is demonstrated in [63]. In case the TV model is not used, the graph construction becomes more complex and the weight of the constraint edges changes. For more details on this point see [63].

Constructing the Ishikawa graph as shown and finding the cut with a minimum cost on this particular graph allows to find the exact (global) optimum solution for the multichannel phase unwrapping problem (6.5). The main disadvantage

of Ishikawa method is related to the memory occupation. In fact, the algorithm stores the whole graph and then performs the cut. This can be a problem when the size of the image increases.

6.4.2 Alpha-expansion algorithm

Alpha-expansion (α -expansion) algorithm has been proposed by Boykov et al. [61] to solve the non-exact optimization problem. This approach is based on the concept of move space. A move is defined as a change of an image configuration f to a new configuration f' . Different types of move exist and in this section we will focus on the expansion move type. An α -expansion move allows any set of image pixels to change their current label into α . It finds a new configuration \hat{f} that minimizes the energy E over all labels f' within this move. This technique is based on computing a labeling corresponding to a minimum cut on a particular graph $\mathcal{G}_\alpha = (\mathcal{V}_\alpha, \mathcal{E}_\alpha)$. The structure of the graph is determined by the current labeling f and the label α . The graph contains S nodes (S is the size of the image) denoted by $\{v_s; s = 1..S\}$ plus two special nodes r and t (the source and the sink). The set of edges \mathcal{E}_α contains two families of edges: data edges and interaction edges. The data edges are related to the multichannel likelihood function (6.2), while the interaction edges are related to the a priori energy function (6.4).

An assumption is made related to the a priori energy function called *submodularity*. In the case of α -expansion move, this assumption implies that the a priori energy function (6.4) has to be a metric to be graph representable, i.e. it has to satisfy a triangular inequality [61]. It is easy to prove that TV function is a metric.

For the construction of the graph, two possible representations exist. The first one is given by Boykov et al. [61] and the second one is given by Kolmogorov et al. [67]. The one given by [61] is used for our purposes. The considered graph, in one dimension, is shown in figure 6.4, in the case of three pixels. The graph is made of 3 vertexes (the pixels s , $s - 1$ and $s + 1$) plus two special vertexes (α , $\bar{\alpha}$ representing respectively the new proposed label and the current labels). In addition, for each pair of neighboring pixels such that their current labels are different, an auxiliary vertex is created. In the case shown in figure, we suppose

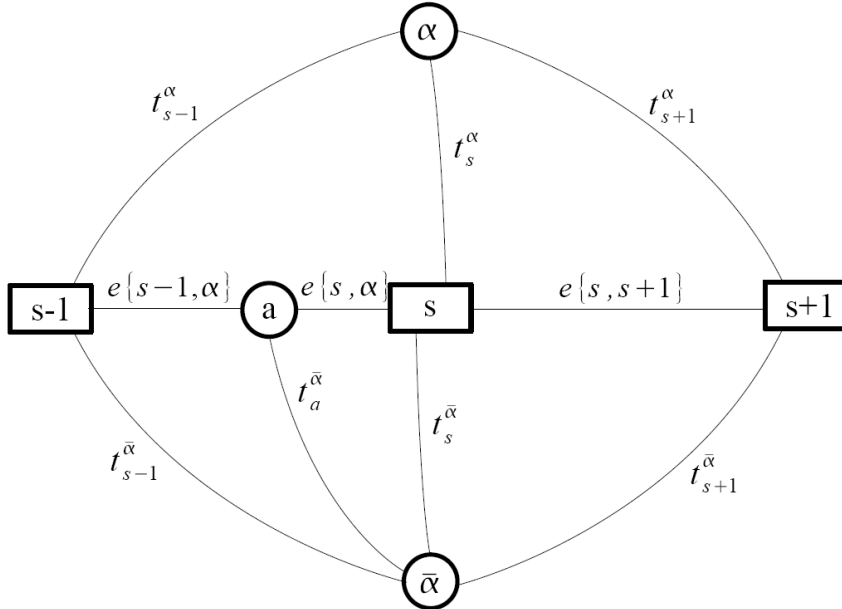


Figure 6.4: α -expansion graph - Representation of the edges for a vertex

that the pixels $s - 1$ and s have different labels ($f_{s-1} \neq f_s$), while pixels s and $s + 1$ have the same labels ($f_s = f_{s+1}$). An auxiliary vertex a is created between the first two pixels (vertex a). Each pixel is connected to the special vertex α with the edges t_s^α and to the vertex $\bar{\alpha}$ with the edge $t_s^{\bar{\alpha}}$. The pixels s and $s + 1$ are connected by an edge $e\{s, s + 1\}$, while for the pixels $s - 1$ and s a triplet of edges is created ($e\{s - 1, a\}, e\{a, s\}, t_a^{\bar{\alpha}}$). In order to be representative of the multichannel phase unwrapping problem, we set the edges as in table 6.1. In table 6.1 $V_{s,s+1}(y, x)$ is equal to the a priori energy function $U(\mathbf{h}, \theta)$, calculated using the label x for the pixel s and the label y for the pixel $s + 1$. We remind that the expression of the data energy is provided by (6.2) and of the a priori energy is provided by (6.4). Constructing the α -expansion graph as shown and finding the cut with the minimum cut on this particular graph, iteratively for all the possible values of the label α allows to find a local optimum within a good quality for the multichannel phase unwrapping optimization problem (6.5). It is proved in [61] that the expansion move algorithm produces a labeling f such that the provided local solution is within a constant factor from the optimal one. Differently from Ishikawa algorithm, α -expansion does not have any problem related to memory

edge	weight	for
$t_s^{\bar{\alpha}}$	∞	$f_s = \alpha$
$t_s^{\bar{\alpha}}$	$F(\Phi(s) \bar{\alpha})$	$f_s \neq \alpha$
t_s^{α}	$F(\Phi(s) \alpha)$	all s
$e\{a, s\}$	$V_{s,s+1}(\alpha, f_s)$	$f_s \neq f_{s+1}$
$e\{s, a\}$	$V_{s,s+1}(f_s, \alpha)$	$f_s \neq f_{s+1}$
$t_{\alpha}^{\bar{\alpha}}$	$V_{s,s+1}(f_s, f_{s+1})$	$f_s \neq f_{s+1}$
$e\{s, s+1\}$	$V_{s,s+1}(f_s, \alpha)$	$f_s = f_{s+1}$

Table 6.1: α -expansion edge weights

capacity. For this reason α -expansion can be used in case of large images. Note that in this case, also the LGMRF model for the a priori energy could be used. This would not make the graph more complicated as stated for Ishikawa case. Anyway, using the LGMRF, the algorithm would suffer again of the time limits underlined before.

6.4.3 Hyperparameter estimation

In both cases, either using with Ishikawa or α -expansion, the hyperparameter θ has to be estimated. Its estimation is essential to obtain a good reconstruction. The hyperparameter θ depends both on data energy and on a priori energy term. Different automatic hyperparameter estimation methods exist and have been mainly developed for the restoration problem [68]. One of the possible methods to perform the hyperparameter estimation is the analysis of the so called *L - curve* [69]. This curve is the graphical representation of the regularization energy term respect to the likelihood energy term. The corner of the curve corresponds to a good trade off between under regularization (where the regularization term can be largely improved with minor likelihood modification) and over regularization proposed (where the regularization term can no longer be improved whatever the likelihood price). To find automatically the corner of the *L - curve*, we have used the triangular method described in [70] providing good results in limited time.

6.5 Simulated Experiments

In order to prove the effectiveness of the proposed method, some results are shown obtained both on simulated and real data (see section 6.6). The results presented were obtained with MATLAB coding (max-flow algorithm is implemented in C++ [62]), using a Pc Core 2 duo 2.66GHz with 2G memory.

For the simulated data two simulation have been conducted. In the first one, the effectiveness of the proposed approach concerning both the model and the optimization algorithms based on graph-cuts is shown. In the second one, an analysis of which kind of optimization algorithm (exact or non-exact) has to be used for different phases that have to be unwrapped has been conducted.

In the first experiment, we considered a height profile (64×64 pixels) with a maximum height of $140m$ exhibiting both smooth and discontinuous areas (Fig. 6.5(a)). We used two frequencies (5 GHz and 9 GHz) to generate interferograms and we added interferometric noise with a coherence of $\gamma = 0.5$. In figure 6.5(b), we show the 5 GHz noisy interferogram. For each working frequency we generated 4 azimuth looks allowing to generate a total of $N = 8$ independent interferograms. It is important to note that the profile is ambiguous for both the working frequencies. In fact, there are phase jumps of about 1.3π at 5 GHz and 2.4π at 9 GHz which violates the Itoh condition (section 3.3.3), i.e. phase differences between neighboring pixels are bigger than π . For this reason, a classical single frequency phase unwrapping method like Branch Cuts (section 3.3.3) would fail. In figures 6.5(c) and 6.5(d) we show the results obtained respectively with the ICM-LGMRF multichannel approach and with the GC-TV approach. We can easily note that both reconstruction profiles appear very similar to the reference one (figure 6.5(a)). The normalized reconstruction square error is equal to 2.05×10^{-2} in the first case and 1.93×10^{-2} in the second.

It is very interesting to see, looking at the square object at the right bottom of figure 6.5(d), how the TV model is well adapted for this kind of profile which simulates a typical urban building as we will see next in real data experiments.

Although reconstruction results are mainly the same using the previous and new approach, their performances in term of execution time are not comparable. In fact, while the new algorithm takes less than one minute, the classical one is 10

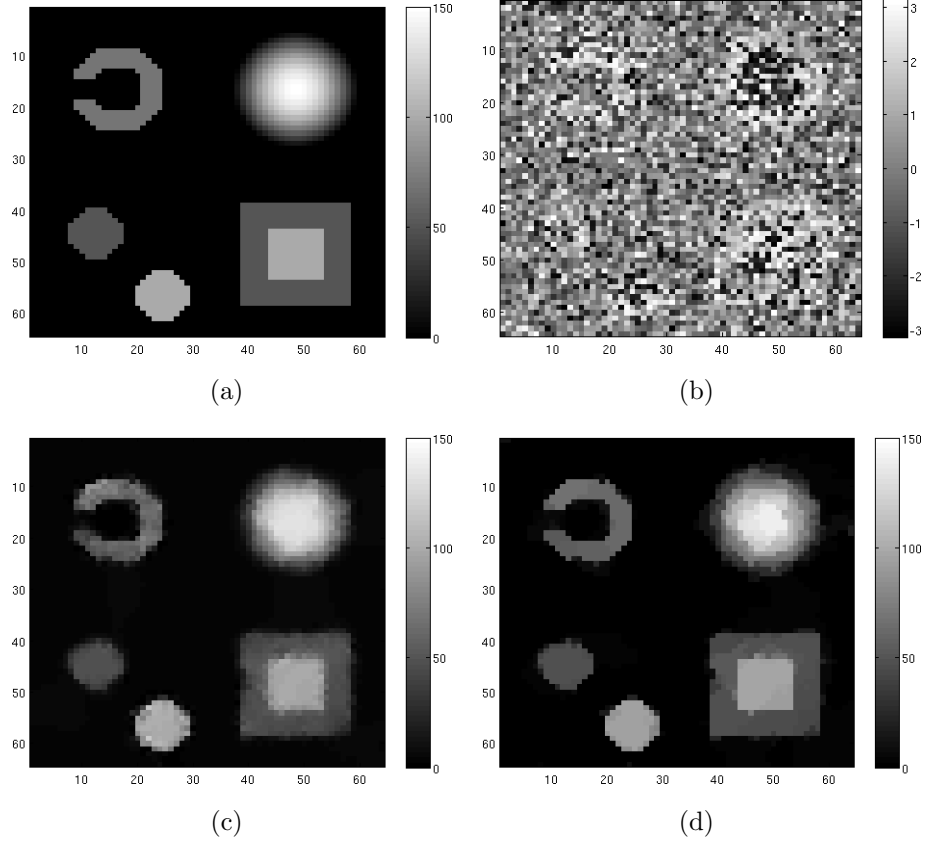


Figure 6.5: (a) Original profile, (b) noisy interferogram at 5 GHz with $\gamma = 0.5$, (c) reconstruction with the multichannel ICM-LGMRF approach, (d) reconstruction with the multichannel GC-TV approach.

times slower (60 seconds for GC-TV and 600 seconds for ICM-LGMRF).

Another point has to be remarked. Since the optimization algorithm based on graph-cut used in this experiment is Ishikawa algorithm, we are sure to obtain the global optimum. Whereas, using ICM in the optimization step we still risk of being trapped in local minima solution. Note that in this case, due to the energy function used in this experiment, also α -expansion algorithm provided the global solution.

In the second experiment, we compare the two graph-cut based optimization algorithms: Ishikawa and α -expansion. The analysis will be based on two characteristics of the wrapped interferogram: size of interferogram and the phase rate.

6. FAST MCPHU

First we fix the phase rate and we consider different interferogram sizes S . We use the reference profile of the first experiment figure 6.5(a) with the same system parameters (coherence, frequencies and channels) and a fixed value of the hyperparameter θ . The results given by the two optimization algorithms are the same (global solution is provided). The computational times (in seconds) are shown in table 6.2. We note two important aspects: first we can see that α -expansion is faster than Ishikawa in all the considered cases. Secondly there exist a thresholds after which using Ishikawa algorithm is impossible. This is due to the memory problem intrinsic in Ishikawa algorithm. Clearly, the threshold depends on the computer architecture. Then, the size of the images is fixed and different inter-

	64×64	128×128	192×192	256×256
TV+ α -expansion	5	22	57	92
TV+Ishikawa	27	114	259	not completed

Table 6.2: Comparison between α -expansion and Ishikawa

ferogram phase rates are considered. In this case, the analysis will be based on the Gaussian profile (160×160 pixels) of figure figure 6.6(a). The maximum height of the profile is 90 rad. The system parameters are the same as for the first experiment. In figure 6.6(b), the 5 GHz noisy interferogram is shown. Due to the large phase rate, large enough to produce aliasing, the energy function presents many local minima that are very near to the global one. This does not allow α -expansion algorithm to reach the global solution. In this case, Ishikawa is the only one adapted and provides the best solution for the multichannel phase unwrapping problem. Figures 6.6(c) and 6.6(d) represent respectively the reconstruction obtained using α -expansion and Ishikawa algorithms. As stated before, the α -expansion fails to retrieve the correct solution, while using Ishikawa we obtain a normalized square reconstruction error of 9.4×10^{-4} . We also note that α -expansion algorithm could be used in this case if additional channels are available. In fact, using more channels, the energy function becomes smoother, the local minima farther to the global one, allowing α -expansion to provide the optimal solution.

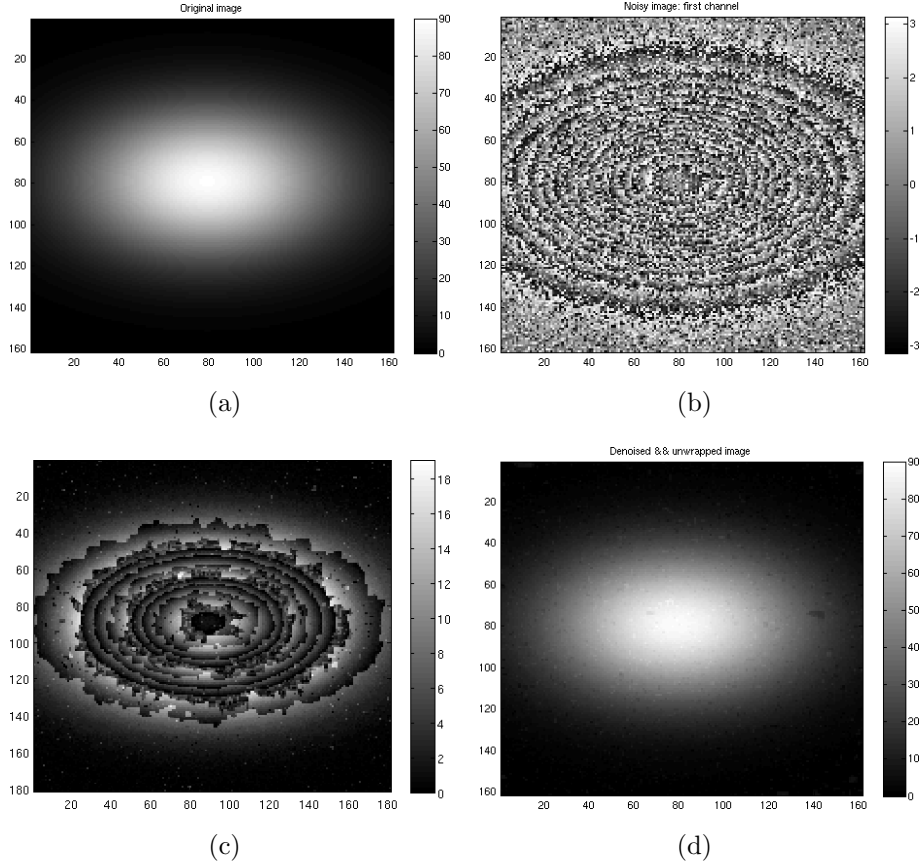


Figure 6.6: (a) Original Gaussian profile, (b) noisy interferogram at 5 GHz with $\gamma = 0.5$, (c) reconstruction with α -expansion algorithm, (d) reconstruction with Ishikawa algorithm.

6.6 Real Data Experiments

We tested our new proposed algorithm on real data set of an urban scenario. We used a set of 8 L-Band E-SAR interferograms (2 interferograms for each of the four polarization) acquired on the city of Dresden. The smallest baseline is of about 10m and the biggest is of about 40m. We applied our new proposed approach based on Ishikawa algorithm and TV model. Even if the height of the building in the first interferogram is less than the ambiguous height ($h_{amb} \in \{37.5m, 11.1m\}$) figure 6.7(a), due to the presence of noise, the Itoh condition is violated in some

6. FAST MCPHU

areas. So, if we apply a classic single channel approach, the solution will show some errors in that areas (figure 6.7(b)). Using our approach, it is possible to retrieve the correct height of the building also in these areas. Note that the height of the roof building is almost constant figure 6.7(c). Moreover, it is possible to note the goodness of the TV model. Thanks to this model, the contours of buildings are recovered very efficiently (figure 6.7(d)).

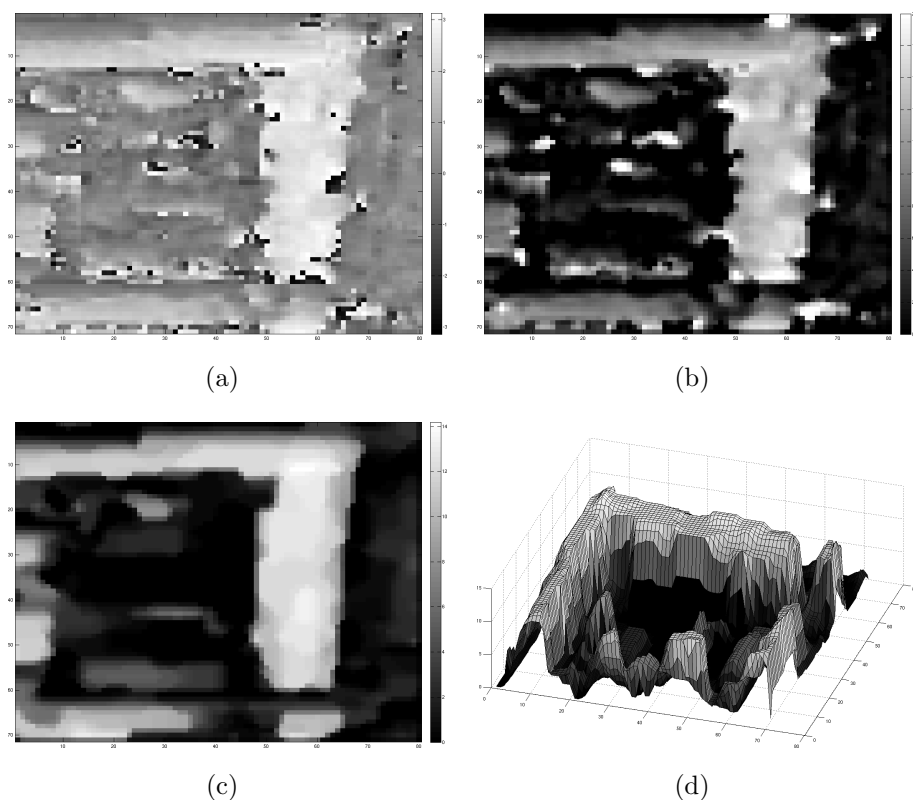


Figure 6.7: (a) Noisy interferogram (smallest baseline), (b) reconstruction with branch-cut algorithm, (c) reconstruction with the proposed multichannel approach, (d) 3D view of the reconstruction with the proposed multichannel approach.

6.7 Conclusions

In this chapter, the development of a new fast and robust multichannel MAP phase unwrapping methodology has been shown. The proposed new method over-

comes the limits that in some cases the multichannel phase unwrapping method based on ICM and LGMRF can presents. The main limits are due to the computational time, which can be in some case large due the estimation of the local hyperparameters. Moreover, the ICM does not give the guarantee to reach the the global optimum. For these reasons a new algorithm is developed. In order to reduce the computational time, a non local energy is used to model the a priori energy. Among the possible non local energy, the Total Variation model has been chosen, since it does not penalize discontinuities in the image while simultaneously not penalizing smooth functions either. In order to reach the global optimum in the minimization step, graph cut based algorithms are used. Two graph-cut based optimization algorithms have been used: Ishikawa algorithm that provides the global optimum and α -expansion algorithm that provides a local optimum within a good quality. The proposed algorithm overcomes the problem of excessively computational times that characterizes the ICM-LGMRF approach. Moreover, with Ishikawa optimization algorithm, we are able to reach the exact energy optimum. The approaches have been tested on simulated data showing good results both in term of reconstruction error and computational time. The approach have been tested also on a real data set, showing a good reconstruction accuracy.

6. FAST MCPHU

7

Further Works

7.1 Introduction

In this chapter further conducted works are presented. In the first part some preliminary studies on Multichannel SAR tomography are presented. Then some preliminary studies on Ground Based SAR are discussed.

7.2 Multichannel SAR Tomography

In chapter 5 the phase unwrapping in urban scenarios has been investigated, underlying the importance of developing a procedure able to correctly manage layover areas. The procedure proposed in section 5.4, based on Map multichannel reconstruction together with auxiliary external data, has proved to guarantee good results under the hypothesis of considering among the three contributes which collapse in each resolution cell only the dominant one. Due to this hypothesis a generalization to all real situations is not ensured. The proposed approach allows to well localize and reconstruct the contribute related to the vertical wall. The geocoded obtained results highlight a drawback: a loosing of the signal components relatives to regions of weak echo response is implicit, producing no signal areas in final DEM (ground and roof area involved in the layover).

Preliminary studies on another solution to handle layover have been conducted. These studies are related to SAR Tomography [71], [72].

7. FURTHER WORKS

SAR tomography consists of trying to separate and distinctly estimate each of the three complex contribute which collapse in layover pixels, overcoming the limits of the previous approach. To do this, a classical statistical estimation techniques can be adopted, in particular Maximum Likelihood Estimation (MLE). To avoid the limitations due to the presence of ambiguous solutions, which are intrinsic in the single interferogram configuration, and to improve the performances of the estimator, the multichannel configurations (either multifrequency or multibaseline) is used allowing to exploit multiple independent data.

Consider two SAR complex images, u_1 and u_2 , acquired with a certain baseline b . The two complex signals can be modelled as:

$$\begin{aligned} u_1 &= s_{1a} + s_{1b} + s_{1c} + n_1 = u_{1R} + ju_{1I} \\ u_2 &= s_{2a} + s_{2b} + s_{2c} + n_2 = u_{2R} + ju_{2I} \end{aligned} \quad (7.1)$$

where s_{1x} and s_{2x} for $x = a, b, c$ are the complex signals acquired by the two antennas related to the three scatterers a, b, c (see figure 7.1 for the geometry of the problem) which collapse in the resolution cell (r_{abc} of figure 7.1) relative to the layover area. Suppose that the three signals come from three different

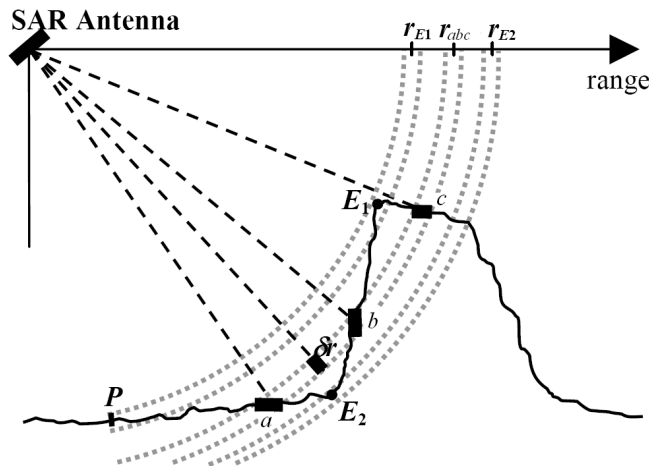


Figure 7.1: Layover Geometry - The three scatterers collapse in the same resolution cell

quotas h_a , h_b and h_c , (see figure 7.1). n_1 and n_2 represent the thermal noise at

7.2 Multichannel SAR Tomography

the receivers. s_{1a} and s_{2a} , s_{1b} and s_{2b} , s_{1c} and s_{2c} will exhibit different phases related to the three different quotas (3.24):

$$\begin{aligned}
 \phi_a &= \left\langle \frac{4\pi B}{\lambda R_1 \sin \theta} \cos(\theta - \alpha) h_a \right\rangle_{2\pi} \\
 \phi_b &= \left\langle \frac{4\pi B}{\lambda R_1 \sin \theta} \cos(\theta - \alpha) h_b \right\rangle_{2\pi} \\
 \phi_c &= \left\langle \frac{4\pi B}{\lambda R_1 \sin \theta} \cos(\theta - \alpha) h_c \right\rangle_{2\pi}
 \end{aligned} \tag{7.2}$$

All the parameters have been previously defined (section 3.3).

We assume:

$$\begin{aligned}
 s_{1a} &= a_1, & s_{2a} &= a_2 e^{j\phi_a} \\
 s_{1b} &= b_1, & s_{2b} &= b_2 e^{j\phi_b} \\
 s_{1c} &= c_1, & s_{2c} &= c_2 e^{j\phi_c}
 \end{aligned} \tag{7.3}$$

a_1 and a_2 are assumed Gaussian complex signals with zero mean, equal variance $2\sigma_a^2$, real and imaginary part uncorrelated. The same assumption is made for b_1 and b_2 , with variance $2\sigma_b^2$, and for c_1 and c_2 , with variance $2\sigma_c^2$. Moreover we assume signals scattered by a , b and c of figure 7.1 independent from each other and from the noise. The expression of the coherence γ between two SAR complex images u_1 and u_2 is given by:

$$\begin{aligned}
 \gamma &= \frac{E[u_1 u_2^*]}{\sqrt{E[|u_1|^2] E[|u_2|^2]}} = \\
 &= \frac{\gamma_a 2\sigma_a^2 e^{-j\phi_a} + \gamma_b 2\sigma_b^2 e^{-j\phi_b} + \gamma_c 2\sigma_c^2 e^{-j\phi_c}}{2(\sigma_a^2 + \sigma_b^2 + \sigma_c^2 + \sigma_n^2)} \left(1 - \frac{B}{B_c}\right)
 \end{aligned} \tag{7.4}$$

where γ_a , γ_b and γ_c are the coherence of each scatterer echo received by the first antenna and the same scatterer echo received by the second antenna without considering the interferometric phase difference between the two signals (i.e. related to the height) and the decorrelation due to distance between acquisition. The term $\left(1 - \frac{B}{B_c}\right)$, where B_c is the critical baseline, takes into account the decorrelation due to baseline length (3.3.2.1).

The two processes $u_1 = u_{1R} + ju_{1I}$ and $u_2 = u_{2R} + ju_{2I}$ are still Gaussian, have

7. FURTHER WORKS

zero mean and a covariance matrix Σ , given by:

$$\Sigma = (\sigma_a^2 + \sigma_b^2 + \sigma_c^2 + \sigma_n^2) \begin{bmatrix} 1 & 0 & \text{Re}(\gamma) & -\text{Im}(\gamma) \\ 0 & 1 & \text{Im}(\gamma) & -\text{Re}(\gamma) \\ \text{Re}(\gamma) & \text{Im}(\gamma) & 1 & 0 \\ -\text{Im}(\gamma) & \text{Re}(\gamma) & 0 & 1 \end{bmatrix}$$

. Then the joint probability density function of $\mathbf{w} = [u_{1R}, u_{1I}, u_{2R}, u_{2I}]^T$, is Gaussian:

$$P(\mathbf{w}) = \frac{1}{2\pi^2 |\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \mathbf{w}^T \Sigma^{-1} \mathbf{w}\right) \quad (7.5)$$

The unknown parameters are $\vartheta = [\sigma_a^2, \sigma_b^2, \sigma_c^2, h_a, h_b, h_c]^T$; once the data are observed \mathbf{w} , the likelihood function is obtained. By taking its logarithmic we obtain:

$$L(\vartheta) = -\log\left(2\pi^2 |\Sigma(\vartheta)|^{\frac{1}{2}}\right) - \frac{1}{2} \mathbf{w}^T \Sigma^{-1} \mathbf{w} \quad (7.6)$$

Using the multichannel configuration $N + 1$ different and independent images (u_1, u_2, \dots, u_N) are obtained using N independent channels. In this configuration the log likelihood function has the same form as (7.6) but with a covariance matrix $2(N + 1) \times 2(N + 1)$ dimensional.

Future works on this subject will regard an analysis of the performances of an unbiased estimators, using the presented model and the implementation of a Maximum Likelihood Estimator in order to retrieve the unknown parameters. Moreover some considerations on the chosen model have to be done in case of different scenarios and different scattering mechanisms.

7.3 Ground Based SAR

Recently a new terrain monitoring technique has been developed based on the use of a Ground Based Synthetic Aperture Radar (GB-SAR) Interferometer [73]. As satellite sensors, GB-SAR acquires complex images of the observed area, where phase is related to the distance from the sensor to the scene. Interferometry exploits this information to recover topography profile and to estimate terrain movements.

In the Interferometric configuration the GB-SAR can monitor deformation of unstable areas having an extension of a few squares kilometres, at time intervals

ranging from few minutes up to several days. The continuous monitoring allows understanding the spatial distribution of deformation and, consequently, to predict the dynamics of the observed phenomenon [74]. Real time capabilities and flexibility, the main features of GB-SAR interferometer, are the main characteristics required to an instrument for the landslide monitoring in emergency cases. Consider the GB-SAR system shown in 7.2 in the orthogonal plane respect the rail along that the synthetic aperture is performed, coinciding with the x axis, where M and S are respectively the master and slave radar positions and $P(x, y, z)$ is the observed target position. Master and slaves antennas are separated by a

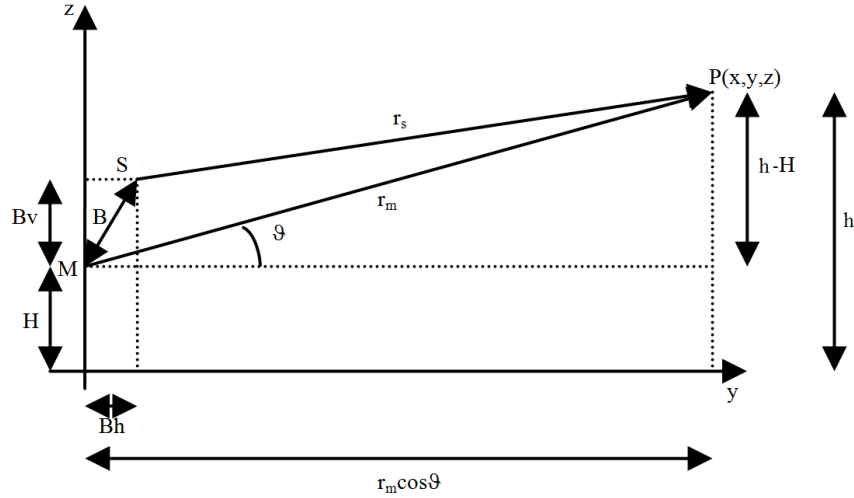


Figure 7.2: Interferometric GB-SAR system - The scatterer point is by the antenna M and by the antenna S , separated by a baseline B

baseline B in which B_v and B_h are the vertical and horizontal components. H represents height of the master antenna respect to the reference plane and r_m is the distance between the master antenna and the target point, r_s is the distance between the slave antenna and the target point, ϑ represents the radar look angle. We assume that the plane passing between observed target and radar is parallel to yz plane, or equivalently the tilt angle between the radar-to-target direction and the orthogonal direction respect to the rail is zero.

Under this assumption, the interferometric phase is given by:

$$\phi = \left\langle \frac{4\pi}{\lambda} (r_m - r_s) \right\rangle_{2\pi} \quad (7.7)$$

7. FURTHER WORKS

The nominal phase ϕ^0 , (the phase in absence of noise) can be related to the topographic height h above the reference plane through the relationship [73]:

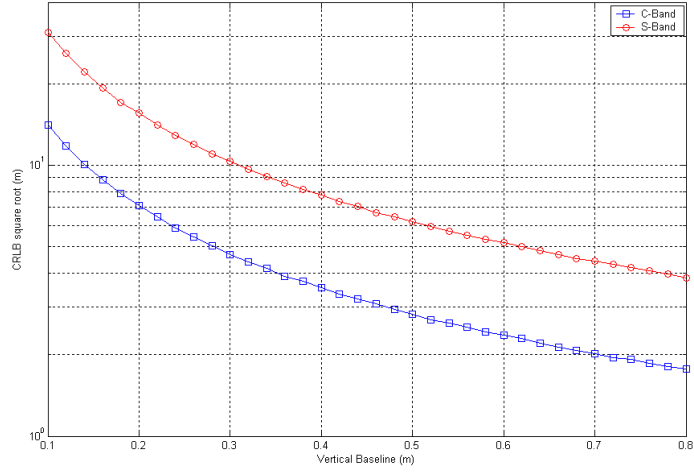
$$\phi^0 = \left\langle \frac{4\pi}{\lambda} \left[B_v \left[\left(\frac{h-H}{r_m} \right) - \frac{B_v}{2r_m} \right] + B_h \left[\sqrt{1 - \left(\frac{h-H}{r_m} \right)^2} - \frac{B_v}{2r_m} \right] \right] \right\rangle_{2\pi} \quad (7.8)$$

The nominal phase differs from the measured phase ϕ , since the measured phase is corrupted by the interferometric noise. The probability density function of ϕ , given the nominal phase ϕ^0 is the same as conventional interferometric SAR systems (3.17).

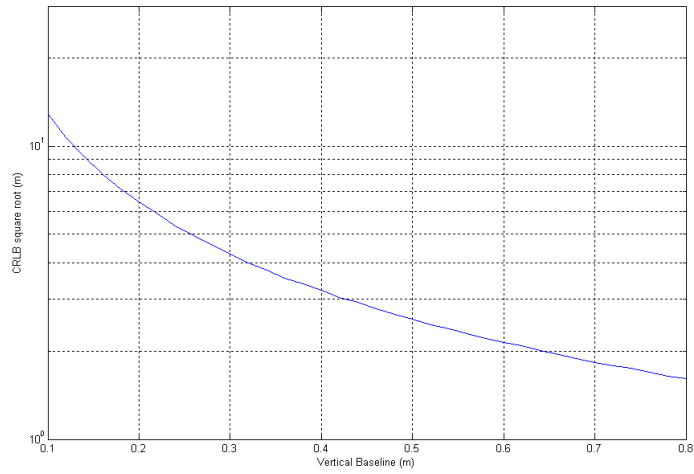
Like for conventional Interferometric SAR systems, also in GB-SAR interferometry, in order to retrieve the height h , the PhU problem has to be solved. A statistical multichannel approach is an effective method also in this case. The derivation of the multichannel likelihood function is the same as in conventional SAR Interferometry (see section 3.4). The only difference is given by the relation between nominal phase ϕ^0 and height h .

Some interesting studies have been conducted on the achievable performances of the ML estimator for the height reconstruction problem in GB-SAR Interferometry. In particular the performances analysis has been conducted evaluating the CRLB. As previously said (see section 2.3), CRLB is an important instrument that can be used to evaluate the performances of an estimation algorithm. We analyze how typical geometric parameters of GB-SAR system can influence CRLBs.

We considered two available frequency bands, C-band and S-band, and a baseline B such that the horizontal baseline B_h is set to zero. Figure 7.3(a) shows the CRLBs square root as function of vertical baseline length B_v : specifically square-marked line refers to estimation using C-band interferogram, while circle-marked line refers to S-Band interferogram. We note that better performance are given by C-Band and also that baseline increasing improves estimation quality. Figure 7.3(b) shows CRLB square root in case of estimation using jointly C and S Bands interferograms: comparing with previous figure, we derive that performance mostly depends on the lower wavelength. However, higher wavelengths give greater robustness to ML estimation approach. Then we evaluated how coherence influences performance: in figure 7.4 CRLBs square root using



(a)



(b)

Figure 7.3: (a) Behaviour of CRLBs square root for increasing vertical baseline length, in C-band and S-band, (b) Behaviour of CRLBs square root for increasing vertical baseline length, using jointly C-band and S-band

jointly C and S Bands data, for increasing vertical baseline length are displayed, considering coherence values varying from $|\gamma| = 0.75$ to $|\gamma| = 0.95$. It has to be

7. FURTHER WORKS

noted that even for high coherence value, small variations of the vertical baseline length cause strong variations in CRLB. .

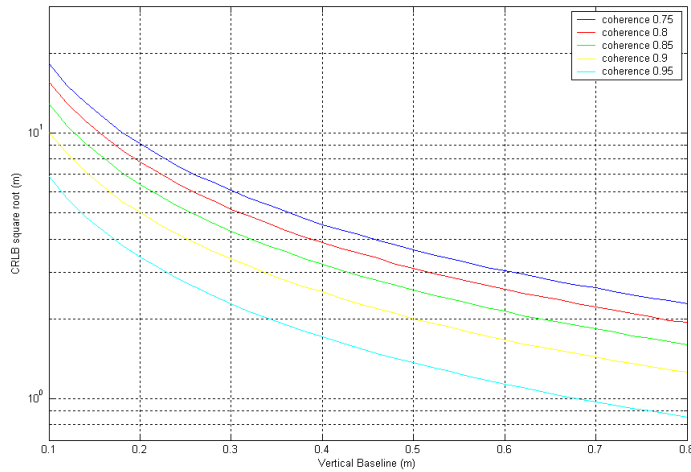


Figure 7.4: GBSAR Interferometry: CRLBs vs coherence - Behaviour of CRLBs square root for increasing vertical baseline length, considering different values of the coherence $|\gamma|$

The next step for this study is to apply the reconstruction procedure to simulated and real data. Moreover, the MAP approach will be implemented in order to improve height reconstruction accuracy.

8

Conclusion

In this thesis the multichannel interferometric systems have been discussed and analysed. Different innovative research contributes in the field of interest have been presented. First, a ML based technique to retrieve the phase offsets that affect multichannel interferometric data has been presented. The interferograms are, in fact, known except for a not known slowly varying phase, that arise from the contribution of several factors. If the phase offsets estimation is not performed, it is not possible to restore the correct relation between the available interferograms. The correct relation between the interferograms is needed by the MAP or ML multichannel phase unwrapping in order to retrieve the height of the considered scene.

The second interesting contribute provided in this thesis is related to the edge detection in urban areas. The proposed idea is to use multichannel interferometric data in order to retrieve buildings edges in urban areas. The approach is based on the estimation of some hyperparameters, used to model the unknown height of the pixels image. The method overcomes the traditional problems typical of SAR edge detectors.

Then a new technique in order to improve the height reconstruction in urban areas, affected by the layover phenomena, has been presented. This technique is based on the use of multichannel InSAR data, in combination with auxiliary data. The auxiliary data can be provided by the shadowing areas in SAR intensity images or by an optical image of the same area. The auxiliary data are used to refine the hyperparameters map, on which the MAP multichannel height

8. CONCLUSION

reconstruction is based. This allows to improve the height reconstruction, in particular on the phase discontinuities that the layover implies.

The last innovative aspect presented in the thesis is a new fast multichannel phase unwrapping method. The proposed method is based on Total Variation Energy Model and Graph cuts optimization algorithm. Towards other classic phase unwrapping techniques, thanks to the multichannel approach the method is able to restore the solution in many critical cases. Moreover, the method overcomes some limits that can arise using other multichannel phase unwrapping techniques. In particular, the proposed method is at the same time not time demanding and robust, since it is able to guarantee the global optimum solution for the considered problem.

Finally a preliminary study on another solution for the height reconstruction in layover areas, and a preliminary study on GB-SAR multichannel interferometric systems have been presented.

References

- [1] S. Z. Li, *Markov random field modeling in image analysis*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2001.
- [2] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.
- [3] C. W. Therrien, *Decision, Estimation and Classification: an Introduction to Pattern Recognition and Related Topics*. New York, Wiley, 1989.
- [4] A. N. Tikhonov and V. A. Arsenin, *Solutions of Ill-posed Problems*. Washington Winston and sons, 1977.
- [5] S. Geman and D. Geman, “Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, November 1984.
- [6] J. Besag, “Spatial interaction and the statistical analysis of lattice systems,” *Journal of the Royal Statistical Society*, vol. B-36, no. 2, pp. 192–236, 1974.
- [7] J. M. Hammersley and P. Clifford, “Markov field on finite graphs and lattices,” *unpublished*, 1971.
- [8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, “Equations of state calculations by fast computational machine,” *Journal of Chemical Physics*, no. 21, pp. 1087–1092.
- [9] B. Chalmond, “An iterative gibbsian technique for reconstruction of m -ary images,” *Pattern recognition*, vol. 22(6), pp. 747–761, 1989.

REFERENCES

- [10] S. Saquib, C. Bouman, and K. Sauer, "Ml parameter estimation for markov random fields, with applications to bayesian tomography," *IEEE Transactions on Image Processing*, vol. 7, pp. 1029–1044, July 1998.
- [11] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via em algorithm," *Journal of the Royal Statistical Society*, no. Series B 39, pp. 1–38.
- [12] S. Kirkpatrick, C. D. Gellatt, and V. M. P., "Optimization by simulated annealing," *Science*, vol. 220, pp. 671–680, 1983.
- [13] G. Franceschetti and R. Lanari, *Synthetic Aperture Radar Processing*. CRC Press, 1999.
- [14] R. K. Raney, *Radar Fundamentals: technical Perspective*, vol. 2. Wiley and Sons, 1198.
- [15] K. Tomiyasu, "Tutorial review of synthetic-aperture radar (sar) with applications to imaging of the ocean surface," *Proceedings of the IEEE*, vol. 66, pp. 563–583, May 1978.
- [16] G. Franceschetti, V. Pascazio, and G. Schirinzi, "Iterative homomorphic technique for speckle reduction in synthetic-aperture radar imaging," *J. Opt. Soc. Am. A*, vol. 12, no. 4, pp. 686–694, 1995.
- [17] R. Bamler and P. Hartl, "Syntetic aperture radar interferometry," *Inverse Problem*, vol. 14, pp. R1–R54, August 1998.
- [18] M. Bao, C. Bruning, and W. Alpers, "Simulation of ocean waves imaging by an along-track interferometric synthetic aperture radar," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 35, no. 3, pp. 618–631, 1997.
- [19] A. Budillon, V. Pascazio, and G. Schirinzi, "Estimation of radial velocity of moving targets by along-track interferometric sar systems," *Geoscience and Remote Sensing Letters, IEEE*, vol. 5, no. 3, pp. 349–353, 2008.

REFERENCES

- [20] D. Massonnet and T. Rabaute, "Radar interferometry: Limites and potential," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, pp. 445–464, 1993.
- [21] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Science*, vol. 31, pp. 445–464, 1993.
- [22] D. Just and R. Bamler, "Phase statistics of interferograms with applications to synthetic aperture radar," *Appl. Opt.*, vol. 33, no. 20, pp. 4361–4368, 1994.
- [23] J. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric sar imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 32, no. 5, pp. 1017–1028, 1994.
- [24] D. Ghiglia and M. Pritt, *Two dimensional phase unwrapping theory, algorithms and software*. New York, Wiley, 1998.
- [25] K. Itho, "Analysis of the phase unwrapping problem," *Applied Optic.*, vol. 2, no. 14, 1982.
- [26] S. Madesn, H. A. Zebker, and J. Martin, "Topographic mapping using radar interferometry: Processing techniques," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 31, no. 1, pp. 246–256, 1993.
- [27] T. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *Journal of the Optical Society of America*, vol. 14, no. 10, pp. 2692–2701, 1997.
- [28] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 3, pp. 813–821, 1998.
- [29] J. Dias and J. Leitao, "The z-pi algorithm: a method for interferometric image reconstruction in sar/sas," *IEEE Transactions on Image Processing*, vol. 11, pp. 408–422, 2002.

REFERENCES

- [30] G. Nico, G. Palubinskas, and M. Dactu, “Bayesian approach to phase unwrapping: Theoretical study,” *IEEE Transactions on Signal Processing*, vol. 48, pp. 2545–2556, 2000.
- [31] D. L. Fried, “Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements,” *Journal of the Optical Society of America*, vol. 67, no. 3, pp. 370–375, 1977.
- [32] A. Ferretti, C. Prati, and F. Rocca, “Multibaseline InSAR DEM reconstruction: The wavelet approach,” *Il Nuovo Cimento*, vol. 24, pp. 159–176, 2001.
- [33] G. Fornaro, A. Monti Guarnieri, A. Pauciullo, and F. De-Zan, “Maximum likelihood multi-baseline SAR interferometry,” *IEE Proceedings Radar, Sonar and Navigation*, vol. 153, pp. 279–288, 2006.
- [34] V. Pascazio and G. Schirinzi, “Estimation of terrain elevation by multi-frequency interferometric wide band SAR data,” *IEEE Signal Processing Letters*, vol. 8, pp. 7–9, January 2001.
- [35] M. Eineder and N. Adam, “A maximum-likelihood estimator to simultaneously unwrap, geocode, and fuse SAR interferograms from different viewing geometries into one digital elevation model,” *Appl. Numer. Math.*, vol. 43, no. 4, pp. 359–373, 2002.
- [36] V. Pascazio and G. Schirinzi, “Multifrequency InSAR height reconstruction through maximum likelihood estimation of local planes parameters,” *IEEE Transactions on Image Processing*, vol. 11, pp. 1478–1489, December 2002.
- [37] F. Ulaby, R. Moore, and A. Fung, *Microwave Remote Sensing: Active and Passive: Vol 3*. Artech House, 1986.
- [38] F. Meglio, V. Pascazio, and G. Schirinzi, “Joint statistical distribution of multi-baseline sar interferograms,” *Geoscience and Remote Sensing Symposium, 2006. IGARSS 2006. IEEE International Conference on*, pp. 2549–2552, 2006.

REFERENCES

- [39] G. Ferraiuolo, V. Pascazio, and G. Schirinzi, “Maximum a posteriori estimation of height profiles in InSAR imaging,” *IEEE Geoscience and Remote Sensing Letters*, vol. 1, pp. 66–70, april 2004.
- [40] F. Rocca, “Modeling interferogram stacks,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, pp. 3289–3299, 2007.
- [41] G. Ferraiuolo, F. Meglio, V. Pascazio, and G. Schirinzi, “Dem reconstruction accuracy in multi-channel sar interferometry,” *to be published in Geoscience and Remote Sensing, IEEE Transactions on*.
- [42] G. Ferraioli, G. Ferraiuolo, and V. Pascazio, “Phase-offset estimation in multichannel sar interferometry,” *Geoscience and Remote Sensing Letters, IEEE*, vol. 5, no. 3, pp. 458–462, 2008.
- [43] G. Ferraioli, “Multichannel sar urban edge detection,” *submitted to Geoscience and Remote Sensing, IEEE Transactions on*.
- [44] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. John Wiley and Sons Inc, 1973.
- [45] W. K. Pratt, *Digital Image Processing: PIKS Inside, 3rd Edition*. Wiley-Interscience, 2001.
- [46] A. Bovik, “On detecting edges in speckle imagery,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 36, pp. 1618–1627, Oct 1988.
- [47] C. Oliver, D. Blacknell, and R. White, “Optimum edge detection in sar,” *Radar, Sonar and Navigation, IEE Proceedings -*, vol. 143, pp. 31–40, Feb 1996.
- [48] R. Fjortoft, A. Lopes, P. Marthon, and E. Cubero-Castan, “An optimal multiedge detector for sar image segmentation,” *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 36, pp. 793–802, May 1998.
- [49] R. Cook, I. McConnell, D. Stewart, and C. Oliver, “Segmentation and simulated annealing,” pp. 30–37, 1996.

REFERENCES

- [50] F. Tupin and M. Roux, “Detection of building outlines based on the fusion of sar and optical features,” *PandRS*, vol. 58, pp. 71–82, June 2003.
- [51] F. Baselice, G. Ferraioli, and V. Pascazio, “Dem reconstruction in layover areas from sar and auxiliary input data,” *submitted to Geoscience and Remote Sensing Letters, IEEE*.
- [52] F. Galland, F. Tupin, J.-M. Nicolas, and M. Roux, “Registering of synthetic aperture radar and optical data,” *Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International*, vol. 5, pp. 3513–3516, 2005.
- [53] M. Suk and S. Jung, “A new image segmentation technique based on partition mode test,” *PR*, vol. 16, no. 5, pp. 469–480, 1983.
- [54] G. Ferraioli, A. Shabou, F. Tupin, and V. Pascazio, “Multichannel phase unwrapping with graph-cuts,” *submitted to Geoscience and Remote Sensing Letters, IEEE*.
- [55] J. M. Bioucas-Dias and G. Valadao, “Phase unwrapping via graph cuts,” *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 698–709, 2007.
- [56] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, no. 1-4, pp. 259–268, 1992.
- [57] M. Nikolova, “A variational approach to remove outliers and impulse noise,” *Journal of Mathematical Imaging and Vision*, vol. 20, pp. 99–120, January 2004.
- [58] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, “An iterative regularization method for total variation-based image restoration,” *SIAM Journal on Multiscale Modeling and Simulations*, vol. 4, no. 2, pp. 460–489, 2005.
- [59] L. Denis, F. Tupin, J. Darbon, and M. Sigelle, “Joint filtering of SAR interferometric phase and amplitude data in urban areas by TV minimization,” *IEEE Geoscience and Remote sensing symposium*, 2008.

REFERENCES

- [60] J. Darbon, “Global optimization for first order markov random fields with submodular priors,” in *IWCIA*, pp. 229–237, 2008.
- [61] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [62] V. Kolmogorov and Z. R., “Multi-camera scene reconstruction via graph cuts,” *ECCV*, vol. 3, pp. 82–96, 2002.
- [63] H. Ishikawa, “Exact optimization for Markov random fields with convex priors,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, pp. 1333–1336, October 2003.
- [64] J. Darbon and M. Sigelle, “Image restoration with discrete constrained total variation part i: Fast and exact optimization,” *Journal of Mathematical Image and Vision*, vol. 26, pp. 261–276, December 2006.
- [65] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: theory, algorithms and applications*. Prentice Hall, 1993.
- [66] Y. Boykov and V. Kolmogorov, “An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision,” in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, pp. 359–374, 2001.
- [67] V. Kolmogorov and R. Zabih, “What energy functions can be minimized via graph cuts?,” in *ECCV (3)*, pp. 65–81, 2002.
- [68] N. Galatsanos and A. Katsaggelos, “Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation,” *IEEE Transactions on Image Processing*, vol. 1, pp. 322–336, July 1992.
- [69] P. Hansen and O. D.P., “The use of the L-curve in the regularization of discrete ill-posed problems,” *SIAM J. Sci. Comput.*, vol. 14, no. 6, pp. 1487–1503, 1993.

REFERENCES

- [70] J. Castellanos, S. Gómez, and V. Guerra, “The triangle method for finding the corner of the L-curve,” *Appl. Numer. Math.*, vol. 43, no. 4, pp. 359–373, 2002.
- [71] F. Gini, F. Lombardini, and M. Montanari, “Layover solution in multibaseline sar interferometry,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, pp. 1344–1356, 2002.
- [72] G. Fornaro and F. Serafino, “Imaging of single and double scatterers in urban areas via sar tomography,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, pp. 3497–3505, 2006.
- [73] G. Antonello, N. Casagli, P. Farina, J. Fortuny, D. Leva, G. Nico, A. Sieber, and D. Tarchi, “A ground-based interferometer for the safety monitoring of landslides and structural deformations,” *Geoscience and Remote Sensing Symposium, 2003. IGARSS '03. Proceedings. 2003 IEEE International*, vol. 1, pp. 218–220 vol.1, 2003.
- [74] M. Pieraccini, G. Luzi, and C. Atzeni, “Ground-based interferometric sar for terrain elevation mapping,” *Electronics Letters*, vol. 36, no. 16, pp. 1416–1417, 2000.

Short form curriculum vitae

Giampaolo Ferraioli was born in Lagonegro, on February, 10 1982. He received the BS and MS degree in Telecommunication Engineering (summa cum laude) at the Università degli Studi di Napoli Parthenope in 2003 and 2005, respectively. His main research interests deal with statistical signal processing, Markov Random Field theory and SAR Interferometry. He is an IEEE student member since 2005. He has been visiting student from February to September 2008 at Department TSI of Telecom ParisTech, Paris, France, supervised by Prof Florence Tupin where he worked on Graph Cut optimization algorithms in multichannel InSAR systems.

Peer reviewed papers

- Ferraioli, G. and Ferraiuolo, G. and Pascazio, V. “Phase Offset Estimation in Multi-Channel SAR Interferometry” *IEEE Geoscience and Remote Sensing Letters*, vol. 5, n. 3, pp. 458-462, 2008.
- Baselice, F. and Ferraioli, G. and Pascazio, V. “DEM Reconstruction in Layover Areas from SAR and Auxiliary Input Data” *IEEE Geoscience and Remote Sensing Letters*, in printing

Submitted papers

- Ferraioli, G. and Shabou, A. and Tupin, F. and Pascazio, V. “Multichannel phase unwrapping with Graph-cuts” *IEEE Geoscience and Remote Sensing Letters*.

- Baselice, F. and Budillon, A. and Ferraioli, G. and Pascazio, V. “Layover Solution in SAR Imaging: a Statistical Approach” *IEEE Geoscience and Remote Sensing Letters*.
- Ferraioli, G. ‘Multichannel SAR Urban Edge Detection’ *IEEE Transactions on Geoscience and Remote Sensing*.

Refereed conference papers

- Ferraioli G, Ferraiuolo G, Pascazio V. “Offset Phase Estimation in Multi-Channel InSAR DEM Reconstruction”, IGARSS 2007, Barcelona, Spain, July 23-27, 2007.
- Ferraioli G, Ferraiuolo G, Pascazio V, Schirinzi G, “Phase Offset Estimation for DEM Reconstruction in Multi-Channel SAR Interferometry”’, Urban Remote Sensing Joint Event, 2007, Paris, France, April 11-13, 2007
- Ferraioli G, Baselice F, Pascazio V “Bayesian DEM Reconstruction from SAR and Optical Data”, IGARSS 2008, Boston, USA, July 6-11, 2008