



PARISAR: Patch-based estimation and regularized inversion for multi-baseline SAR interferometry

Giampaolo Ferraioli, Charles-Alban Deledalle, Loïc Denis, Florence Tupin

► To cite this version:

Giampaolo Ferraioli, Charles-Alban Deledalle, Loïc Denis, Florence Tupin. PARISAR: Patch-based estimation and regularized inversion for multi-baseline SAR interferometry. 2017. ujm-01525973v1

HAL Id: ujm-01525973

<https://ujm.hal.science/ujm-01525973v1>

Preprint submitted on 22 May 2017 (v1), last revised 5 Dec 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PARISAR: Patch-based estimation and regularized inversion for multi-baseline SAR interferometry

Giampaolo Ferraioli, Charles-Alban Deledalle, Loic Denis, Florence Tupin

Abstract—Reconstruction of elevation maps from a collection of SAR images obtained in interferometric configuration is a challenging task. Reconstruction methods must overcome two adverse effects: the strong interferometric noise that contaminates the data, and the 2π phase ambiguities. Interferometric noise requires some form of smoothing among pixels of identical height. Phase ambiguities can be solved, up to a point, by combining linkage to the neighbors and a global optimization strategy to prevent from being trapped in local minima. This paper introduces a reconstruction method, PARISAR, that achieves both a resolution-preserving denoising and a robust phase unwrapping by combining non-local denoising methods based on patch similarities and total-variation regularization. The optimization algorithm, based on graph-cuts, identifies the global optimum. We compare PARISAR with several other reconstruction methods both on numerical simulations and satellite images and show a qualitative and quantitative improvement over state-of-the-art reconstruction methods for multi-baseline SAR interferometry.

Index Terms—SAR interferometry, multi-channel InSAR, Non-local means, TV regularization

I. INTRODUCTION

Phase unwrapping (PhU) operation is one of the most challenging tasks when dealing with three dimensional (3D) reconstruction of earth surface based on Interferometric Synthetic Aperture Radar imaging [1]. PhU consists of retrieving the absolute value of the phase, starting from the 2π -wrapped data. Thanks to the widely known relation between the measured interferometric phase and the height of the observed scene [2], it is possible, after the PhU operation has been correctly performed, to recover the height of the observed area.

Several PhU algorithms have been developed in the last twenty years, and they can be classified into two main families: path-following methods and global optimization methods. Path-following PhU algorithms follow a path in the wrapped phase and unwrap each pixel locally. Algorithms from the second family minimize some measure of misfit between the unwrapped solution and wrapped one while promoting unwrapped solutions with few discontinuities. A good review of these algorithms can be found in [3] and [4].

Two difficulties make PhU a non-trivial operation: the first is due to the perturbations of interferometric noise on the

acquired data; the second is the presence of phase differences larger than π between two neighboring pixels, violating the so-called Itoh condition [5]. Such large phase differences arise when neighboring pixels have very different height values (i.e. in presence of discontinuities), or due to (strong) interferometric noise. Most existing algorithms account for the statistics of interferometric noise. The violation of Itoh condition makes the PhU problem ill-posed, thus challenging to solve. Commonly, to regularize the PhU problem and obtain a unique solution, differences between neighboring absolute phases are supposed to be less than π . This hypothesis is satisfied in the case of height profiles without strong discontinuities and high slopes, and for small baseline values [1].

PhU can be applied to more complex scenes with strong discontinuities or steep slopes by increasing the number of interferograms used during the inversion. By correctly combining different available interferograms, it is possible to restore the solution uniqueness without imposing constraints on the phase difference between neighboring pixels [6]. Multiple interferograms, commonly known as multi-channel interferograms, can be obtained in two different ways: using sensors working at different frequencies or using sensors acquiring the scene with different baselines. The latter, multi-baseline interferometry, is the case when the sensor observes the same scene, repeatedly, from slightly different positions, and is commonly the adopted one [7].

In the past years, multi-baseline PhU techniques have been largely investigated [8], [9], [10], [11]. More recently, new multi-baseline algorithms have been proposed. A technique based on the extension of cluster analysis has been proposed in [12]. The reduction of memory requirements when dealing with multiple data is the main aim of [13]. The use of Kalman Filter in case of multiple acquisitions has been investigated in [14]. Finally, multi-baseline interferograms have also been used together with other information to improve reconstruction accuracy in urban areas: in [15] multi-baseline data have been jointly exploited with multi-aspect data while in [16] multi-baseline interferograms have been exploited together with amplitude information.

In order to obtain satisfying results using multi-baseline data, the first step is to correctly combine the available information. An effective way to combine the available multi-channel (i.e., multi-baseline) interferometric data is to exploit statistical estimation methods. These methods propose to exploit the statistical distribution of the acquired data and to implement instruments provided by both classical [17], [18] and Bayesian estimation theory. In particular, for the latter when Markov Random Fields (MRF) theory is used for modeling

G. Ferraioli is with Dipartimento di Scienze e Tecnologie Università degli Studi di Napoli “Parthenope” – Naples, Italy

C.A. Deledalle is with CNRS-Univ. Bordeaux, IMB – Talence, France

L. Denis is with Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d’Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France

F. Tupin is with LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France

Manuscript received XXX; revised XXX

the unknown height profile the so-called Bayesian Markovian estimation framework arises [19], providing very effective results in the multi-channel case [20], [21]. An interesting work based on Bayesian Markovian framework, developed in the single channel case has been recently proposed in [3].

In this paper we propose to exploit contextual information to improve multi-baseline unwrapping. Patch-based approaches, like NL-SAR [22], can efficiently exploit local structural information in the noisy signal to gather similar samples and improve the estimation. To do so, they compare small pieces of information (the patches) and combine the similar ones. These estimators produce results with non-stationary residual variance: in regions where many similar patches are found, the estimate is accurate, while rare configurations are left almost unchanged, *i.e.*, with the strong original interferometric variance. At these locations, an additional smoothing is to be enforced. Moreover, ambiguities due to phase wrapping can often be solved based on local smoothness priors. Markovian prior models of the elevation can be defined in this regard: *total variation* (TV) or truncated quadratic functions lead to smooth elevation while allowing strong discontinuities [16]. These regularization models applied alone suffer some limits like *staircasing* effects affecting flat areas and leading to piecewise constant reconstruction [23]. Following the approach proposed in [24] for image and video, we investigate the combination of both a patch-based approach and TV regularization for elevation estimation in a multi-baseline interferometric framework, exploiting the whole statistical distribution of the interferometric data.

The adopted strategy includes two steps: first a filtering method like NL-SAR provides non-local (*i.e.*, patch-based) estimates of the covariance matrices at each pixel from the multi-channel images available, while preserving at best the resolution; second, height values are regularized using an edge-preserving regularization such as TV. Following the approach proposed in [24], the problem can be formulated as a maximum *a posteriori* estimation with a modified likelihood term. Section II describes the proposed model: the proposed likelihood term introducing patch-based similarity is described, the prior TV term and the global energy to be minimized, as well as the adopted optimization scheme. In section III, an in depth study of the proposed model is provided through experiments on simulated data, while results on real images are presented and discussed in section IV.

II. THE MODEL

At a given pixel i of a multi-channel interferogram with D channels, the D complex amplitudes can be collected in a vector \mathbf{g}_i , called the scattering vector. Under the classical hypothesis of fully developed speckle (Goodman model [25]), the scattering vector \mathbf{g}_i is distributed according to a circular complex Gaussian:

$$p(\mathbf{g}_i | \Sigma_i) = \frac{1}{\pi^D \det(\Sigma_i)} \exp(-\mathbf{g}_i^\dagger \Sigma_i^{-1} \mathbf{g}_i) \quad (1)$$

with \mathbf{g}_i^\dagger the Hermitian transpose of the column vector \mathbf{g}_i . This distribution is parameterized by the $D \times D$ complex covariance

matrix $\Sigma_i = \mathbb{E}[\mathbf{g}_i \mathbf{g}_i^\dagger]$ (\mathbb{E} denoting the expectation) at pixel i . Considering that all the channels have the same radiometry R , and denoting by $s_{a,b} = \mathbb{E}[g_i(a)g_i(b)^*]/R = \gamma_{a,b} \exp(j\psi_{a,b})$ the inter-channel correlation, with coherence $\gamma_{a,b}$ and interferometric phase $\psi_{a,b}$, leads to:

$$\Sigma = R \begin{pmatrix} 1 & s_{1,2} & \cdots & s_{1,D} \\ s_{1,2}^* & 1 & & s_{2,D} \\ \vdots & & \ddots & \vdots \\ s_{1,D}^* & s_{2,D}^* & & 1 \end{pmatrix}. \quad (2)$$

The interferometric phases $\psi_{a,b}$ are related to the height h through a function $f_{a,b}$ that accounts for the interferometric baseline, eventual atmospheric distortions and other calibration parameters [1]:

$$\psi_{a,b} = f_{a,b}(h) = \alpha_{a,b} h = \frac{4\pi B_\perp(a,b)}{\lambda R_0 \sin \theta} h, \quad (3)$$

where λ is the working wavelength, $B_\perp(a,b)$ is the orthogonal baseline between channels a and b , c is the speed of light, R_0 is the distance to the scene, and θ is the view angle.

In multi-baseline interferometry, a first step generally consists of estimating the covariance matrix Σ_i at pixel i by spatial averaging over a square window \mathcal{W}_i centered on i :

$$\hat{\Sigma}_i^{(\text{box})} = \frac{1}{N} \sum_{j \in \mathcal{W}_i} \mathbf{g}_j \mathbf{g}_j^\dagger. \quad (4)$$

N being the number of samples in \mathcal{W}_i . The interferometric phases $\psi_{a,b}$ extracted from this empirical covariance matrix are then inverted, in a second step, to produce an estimate \hat{h} of the height such that $\hat{\psi}_{a,b} \approx f_{a,b}(\hat{h})$ for all channels a and b .

Such an approach suffers from two drawbacks: (i) the first step involves an averaging procedure that degrades the spatial resolution by blurring thin structures, and (ii) the height estimation does not consider estimated heights at neighboring locations, thereby producing very noisy estimates in low coherence regions.

In order to address these drawbacks, we propose to follow a Maximum a Posteriori (MAP) approach. In Bayesian estimation theory, a MAP estimator is computed by minimizing the *a posteriori* energy \mathcal{E} , which is the sum of two terms: the *likelihood term* (aka “data term” \mathcal{D}) and the *a priori term* (aka “regularization” \mathcal{R}). The bias and variance of the estimator are controlled by balancing the relative weight of those two terms. Given the strong fluctuations of point estimates of interferometric phase, we consider in paragraph II-A a generalization of the *likelihood term* to include a form of averaging over similar pixels within an extended neighborhood. The smoothing enforced by the *a priori term* to produce a satisfying estimate has then no need to be as severe as for a point estimate. We discuss the definition of the *a priori term* in paragraph II-B.

A. Weighted likelihood term

The statistical model defined by Eq.(1) leads to the following likelihood term at pixel i (with *const.* a constant term):

$$-\log p(\mathbf{g}_i|\Sigma_i) = \log \det(\Sigma_i) + \mathbf{g}_i^\dagger \Sigma_i^{-1} \mathbf{g}_i + \text{const.} \quad (5)$$

The number of unknowns in Σ_i is larger than the number D of observations in \mathbf{g}_i . Estimation of h alongside of R and $\gamma_{a,b}$ values with a MAP estimator would thus rely on the choice of regularization terms expressed on all these unknowns. Designing such a regularization may be difficult due to the different nature of the unknowns: radiometry, coherence, and height, and their non-linear interaction in the definition of Σ_i in Eq.(2). To circumvent these problems, we choose to replace the likelihood term of covariance matrix Σ_i with a more general expression: the weighted likelihood [26], [27], [28]:

$$\mathcal{D}_i = - \sum_j \omega_{i,j} \log p(\mathbf{g}_j|\Sigma_i) \quad (6)$$

with $\omega_{i,j}$ the weight given to scattering vector \mathbf{g}_j in the estimation at pixel i . In words, the covariance Σ_i is not only required to support the observation at pixel i but also observations at all the pixels j for which the weights $\omega_{i,j}$ are large. Setting the weights $\omega_{i,j}$ to be equal to each other within the square window \mathcal{W}_i centered on i and equal to 0 outside leads to Eq.(4), *i.e.*, the boxcar covariance estimator. Spatially extending the number of observations related to a given covariance matrix Σ_i reduces the need for a regularization since the number of unknowns becomes much fewer than the number of observations. This however comes at a price: by mixing observations from different spatial locations j in the estimation of Σ_i , the spatial resolution is reduced. It is therefore crucial that the weights $\omega_{i,j}$ be carefully chosen so as to include in Eq.(6) only pixels corresponding to the same covariance Σ_i . Designing methods to adaptively compute weights that preserve at best the resolution has been the subject of numerous works, starting with Lee's sigma filter [29] and oriented windows [30] up to more recent patch-based methods, see the review [31]. In the following, we chose to compute the weights using the NL-SAR algorithm [22] since it is very effective at preserving fine structures, and its parameters are tuned in an unsupervised way to adapt to the number of channels D , the sensor, the resolution and the image content. The derivation of our method is however general and independent from the choice of a specific algorithm for computing the weights $\omega_{i,j}$.

We define first the *weighted maximum likelihood estimator* $\hat{\Sigma}_i^{(\text{WML})}$ as the covariance matrix Σ_i that minimizes \mathcal{D}_i .

Proposition 1. *The weighted maximum likelihood estimator is given by the following weighted averaging:*

$$\hat{\Sigma}_i^{(\text{WML})} = \frac{1}{\lambda_i} \sum_j \omega_{i,j} \mathbf{g}_j \mathbf{g}_j^\dagger, \quad (7)$$

with $\lambda_i = \sum_j \omega_{i,j}$.

Proof. The weighted maximum likelihood estimator is defined

by:

$$\begin{aligned} \hat{\Sigma}_i^{(\text{WML})} &= \arg \min_{\Sigma_i} - \sum_j \omega_{i,j} \log p(\mathbf{g}_j|\Sigma_i) \\ &= \arg \min_{\Sigma_i} \sum_j \omega_{i,j} [\log \det(\Sigma_i) + \mathbf{g}_j^\dagger \Sigma_i^{-1} \mathbf{g}_j] \\ &= \arg \min_{\Sigma_i} \lambda_i \log \det(\Sigma_i) + \sum_j \omega_{i,j} \text{tr} \left[\Sigma_i^{-1} \mathbf{g}_j \mathbf{g}_j^\dagger \right] \\ &= \arg \min_{\Sigma_i} \lambda_i \log \det(\Sigma_i) + \text{tr} \left[\Sigma_i^{-1} \left(\sum_j \omega_{i,j} \mathbf{g}_j \mathbf{g}_j^\dagger \right) \right] \end{aligned}$$

The gradient of the objective function with respect to Σ_i is:

$$\lambda_i \Sigma_i^{-\dagger} - \Sigma_i^{-\dagger} \left(\sum_j \omega_{i,j} \mathbf{g}_j \mathbf{g}_j^\dagger \right) \Sigma_i^{-\dagger}.$$

After multiplication from the left and from the right by Σ_i^\dagger , the first order optimality condition (gradient is null) leads to the desired result. \square

The expression of the data term \mathcal{D}_i can be significantly simplified into a single term thanks to the following proposition:

Proposition 2. *The weighted likelihood data term can be written in terms of the weighted likelihood estimate:*

$$\mathcal{D}_i = \lambda_i \left(\log \det(\Sigma_i) + \text{tr} \left[\Sigma_i^{-1} \hat{\Sigma}_i^{(\text{WML})} \right] \right).$$

Proof. The weighted likelihood data term has been defined in equation (6) by:

$$\begin{aligned} \mathcal{D}_i &= - \sum_j \omega_{i,j} \log p(\mathbf{g}_j|\Sigma_i) \\ &= \sum_j \omega_{i,j} [\log \det(\Sigma_i) + \mathbf{g}_j^\dagger \Sigma_i^{-1} \mathbf{g}_j] \\ &= \lambda_i \log \det(\Sigma_i) + \text{tr} \left[\Sigma_i^{-1} \left(\sum_j \omega_{i,j} \mathbf{g}_j \mathbf{g}_j^\dagger \right) \right] \\ &= \lambda_i \log \det(\Sigma_i) + \lambda_i \text{tr} \left[\Sigma_i^{-1} \hat{\Sigma}_i^{(\text{WML})} \right] \end{aligned}$$

Since the data term \mathcal{D}_i will be involved in minimization problems, we dropped the irrelevant additive constant term. \square

Proposition 2 has important practical consequences. While the original definition of the weighted likelihood data term \mathcal{D}_i involved the sum of many terms (typically, several hundred in the context of non-local methods) for a single pixel i , introduction of the weighted maximum likelihood estimate drastically simplifies the expression of \mathcal{D}_i into a single term. This paves the way to a maximum *a posteriori* estimation based on data terms \mathcal{D}_i .

B. Prior term

In urban areas and at meter resolutions, the height is typically constant from one pixel to a neighboring pixel, or varies strongly when the two pixels belong to two different structures, *e.g.*, ground and roof. We therefore select a prior

term that favors piecewise constant images: the total variation defined by

$$\sum_{(i,j)} |h_i - h_j|, \quad (8)$$

where (i, j) indicates a pair of neighboring pixels.

C. A posteriori energy \mathcal{E}

The *a posteriori* energy includes both the data term introduced in paragraph II-A and the regularisation proposed in paragraph II-B:

$$\mathcal{E}(\mathbf{h}) = \sum_i \mathcal{D}_i + \beta \sum_{(i,j)} |h_i - h_j|, \quad (9)$$

where β is a hyper-parameter that balances the relative importance of the fidelity to the observations (enforced by terms \mathcal{D}_i) and the smoothness of the height map \mathbf{h} (enforced by the *a priori*). Beyond this global tuning through parameter β , it is necessary to account for the variable number of neighbors included in the weighted likelihood. Weights $\omega_{i,j}$ indeed vary from one pixel to another. In homogeneous regions, many similar neighbors are identified, thus the weighted maximum likelihood estimate $\hat{\Sigma}_i^{(WML)}$ is reliable. In contrast, in an isolated structure, very few similar neighbors are identified and most weights $\omega_{i,j}$ are (close to) zero, leading to a very noisy estimate $\hat{\Sigma}_i^{(WML)}$. To account for this disparity between estimates, we follow the idea of [24] and set the sum of weights λ_i at pixel i (see eq.(7)) to be inversely proportional to the standard deviation of the estimator:

$$\lambda_i = \sqrt{\hat{L}_i}, \quad (10)$$

with \hat{L}_i the equivalent number of looks corresponding to the weighted neighborhood defined by the weights $\omega_{i,j}$ [22]:

$$\hat{L}_i = \frac{(\sum_j \omega_{i,j})^2}{\sum_j \omega_{i,j}^2}. \quad (11)$$

D. MAP estimation of the height distribution

The height map \mathbf{h} , *i.e.*, the vector of heights for all pixels, can be estimated in the MAP sense by solving the minimization problem:

$$\hat{\mathbf{h}}^{(MAP)} = \arg \min_{\mathbf{h}} \mathcal{E}(\mathbf{h}). \quad (12)$$

Since $\det(\Sigma_i)$ does not depend on h_i , see [32], the $\log \det(\Sigma_i)$ terms can be dropped in the data terms \mathcal{D}_i . The energy minimization problem becomes:

$$\hat{\mathbf{h}}^{(MAP)} = \arg \min_{\mathbf{h}} \sum_i \sqrt{\hat{L}_i} \cdot \text{tr} \left[\Sigma_i^{-1} \hat{\Sigma}_i^{(WML)} \right] + \beta \sum_{(i,j)} |h_i - h_j|. \quad (13)$$

This minimization problem is highly non-convex because of the dependence on \mathbf{h} in the data term through a phase term. However, global minimization can still be performed since the data term is separable (a sum of independent terms over all

pixels) and the regularization is a sum of convex pairwise terms (*i.e.*, involving only pairs of pixels). We use the graph construct of Ishikawa [33] to map the original non-convex problem into a maximum-flow / minimum-cut problem. We discretize the range of height values into H heights, then build a graph with H layers, each layer containing a node for each pixel in the image. Each node is connected to nodes corresponding to the spatial neighbors within each layer, and to the corresponding nodes in the layer immediately above and below. Capacities of the edges are set according to values of the terms in the optimization problem (13).

The height reconstruction algorithm, called PARISAR (PAtch-based estimation and Regularized Inversion for SAR interferometry), is summarized in the following box.

Algorithm: PARISAR Height Reconstruction Algorithm

- 1 Collect the D SLC images \mathbf{g}
 - 2 Estimate $\hat{\Sigma}_i^{(WML)}$ and \hat{L}_i for all pixels i (*e.g.*, with NL-SAR)
 - 3 **for all** h in the discretized range of heights
 - 4 **for all** pixels i
 - 5 Compute the data term $\sqrt{\hat{L}_i} \text{tr} \left[\Sigma_i^{-1} \hat{\Sigma}_i^{(WML)} \right]$
 - 6 Add a node in the graph for pixel i
 - 7 with an edge capacity equal to the data term
 - 8 Add edges between neighboring nodes
 - 9 **end for**
 - 10 Compute the minimum cut on the graph
 - 11 Derive the optimal height map $\hat{\mathbf{h}}^{(MAP)}$ from the minimum cut
-

III. VALIDATION ON NUMERICAL SIMULATIONS

The quantitative and qualitative assessment of the method has been conducted on different test cases. First a quantitative validation is performed using three different simulated test cases: an urban-like scenario, a pattern of squares and a natural scenario, named Ghiglia. The first test case aims at validating the ability of the proposed approach to unwrap and regularize areas characterized by height discontinuities, to correctly handle very low coherence areas (for example shadow areas), and to retrieve small scale structures. The second simulation is designed to analyze the performance of the method for a wide range of configurations (*i.e.*, different coherence and building height values). The last simulation, Ghiglia, is a classical numerical simulation used to assess the behavior of algorithms in a case close to natural height maps.

The datasets are made of three complex images corresponding to different baselines. The adopted system parameters are summarized in Table I. The processing is performed by a high-level code written in Matlab language and based on C/C++ code for NL-SAR filtering and minimum cut computations, on an Intel Core i7 workstation with Linux Debian as operative system.

A. Urban-like profile

The complex data have been simulated starting from the height profile and the coherence map, reported in Figures 1(f) and 1(a). The starting height profile is a typical scenario

TABLE I
NUMERICAL SIMULATIONS PARAMETERS (SIZES GIVEN IN PIXELS).

| Dataset | $\alpha_{a,b}$ | h_{amb} | size |
|---------|--------------------|----------------|------------------|
| Urban | [-0.55 -1 -0.45] | [5.7 3.1 6.9]m | 64×64 |
| Squares | [-0.55 -1 -0.45] | [5.7 3.1 6.9]m | 240×240 |
| Ghiglia | [-0.55 -1.2 -0.65] | [5.7 2.6 4.8]m | 458×157 |

TABLE II
NORMALIZED RECONSTRUCTION SQUARE ERROR

| Dataset | MLNL | MAPNL | MCPU | PUMA | PARISAR |
|---------|------|-------|------|------|---------|
| Urban | 2.79 | 0.61 | 0.76 | – | 0.01 |
| Squares | 3.75 | 1.00 | 3.86 | – | 0.48 |
| Ghiglia | 0.30 | 0.14 | – | 0.01 | 0.002 |

used to assess the capability of multi-channel algorithms, made of structures of different height, characterized by height discontinuities. The height of the buildings are such that the related phases are ambiguous even for the smallest considered baseline (see Table I), making the unwrapping of the profile a difficult task, even in absence of noise. Concerning the coherence map, three different values have been adopted (see the corresponding cells in Table I). The smaller values have been considered in order to simulate a low coherence area and a shadowing-like area. Note that for the considered profile and the adopted coherence the Itoh condition is not satisfied, thus a single-channel unwrapping algorithm can not be adopted.

The generated $D = 3$ interferograms are shown in Figures 1(c), 1(d), 1(e). The mean estimated coherence map, using a simple box-car filter, is shown in Figure 1(b).

In order to compare the results provided by the proposed algorithm PARISAR, different multi-channel methods have been considered: the MLNL approach (using only the likelihood terms \mathcal{D}_i), the MAPNL approach (using Eq.(13) and a sub-optimal minimization procedure based on Iterated Conditional Modes (ICM)), the MCPU proposed in [21].

From the visual inspection of the results, the good performances of the proposed algorithm are evident. While all the other considered techniques fail in estimating the height of some buildings (MCPU), fail in removing the noise (MLNL) or fail in retrieving the details of the image, such as borders, the small structures or shadowing areas (MAPNL), PARISAR is able to correctly solve all the previously reported issues. The image is well regularized, all the structures, with the correct heights, are retrieved. Shadow areas are well reconstructed and the small structure is not flatten or confused with the surrounding ground area. A strong reduction of the variance of height estimation while preserving edges (no blurring phenomenon) is achieved.

The visual analysis is confirmed by the quantitative analysis of the Urban-like profile based on the evaluation of the Normalized Reconstruction Square Error (defined as in [20]) of Table II. and on the Root Mean Square Error, reported in Table III. For both parameters, PARISAR outperforms the other algorithms.

TABLE III
ROOT MEAN SQUARE ERROR

| Dataset | MLNL | MAPNL | MCPU | PUMA | PARISAR |
|---------|-------|-------|------|------|---------|
| Urban | 5.27 | 2.46 | 2.76 | – | 0.35 |
| Squares | 5.20 | 2.69 | 5.28 | – | 1.87 |
| Ghiglia | 17.88 | 12.11 | – | 3.96 | 1.44 |

B. Pattern of squares

This second test aims at analyzing the proposed method for different configurations, in terms of height and coherence values. The starting true profile is made of structures of different heights (Figure 2(f)), while the coherence spans different values from 0 to 1 (Figure 2(a)). By combining the true profile and the coherence map with the different available baselines (system parameters are reported in Table I), three complex images are generated. The effect of this combination is an interferogram characterized by different behaviors: the top left part contains small structures with small noise (high coherence value), while the bottom right corner is characterized by high structures with large noise. The other two quadrants contain small structures with large noise and high structures with small noise.

The dataset is used to test all previously mentioned algorithms. The results are shown in the second row of Figure 2. MLNL and MCPU provide unsatisfying results: the former provides a noisy solution, while the latter over regularize the solution creating artifacts. The best results are achieved in case of MAPNL and PARISAR. Exploiting the non-local estimation both techniques are able to provide effective results in almost all the areas. PARISAR outperforms MAPNL, in terms of noise regularization and correct height retrieval. This is evident from both visual inspection and quantitative analysis, reported in Tables II and III (see the corresponding lines for the Square dataset). The errors that appear in PARISAR reconstruction are mainly in the bottom right corner, as could be expected since this area is characterized by high structures and strong interferometric noise.

C. Ghiglia profile

In order to assess the performances of the algorithm on a natural scenario, a realistic profile generated on the basis of a real digital elevation model of mountainous terrain around Long's, and isolation Peak Colorado, is considered [34]. In the following we will refer to it as Ghiglia profile. The system parameters are reported in Table I. Three different coherence values, [0.70.650.6], are adopted, for the three considered combination of images. Three interferograms are generated. The true profile, the data and the results are reported in Figure 3. The considered profile is not ambiguous: there are no height discontinuities. In this case, the unwrapping task difficulty comes from the fringes that tend to overlap, creating a sort of aliasing. Since the profile is not ambiguous, a single-channel phase unwrapping can be used to unwrap the profile. The PUMA algorithm proposed in [35] is considered.

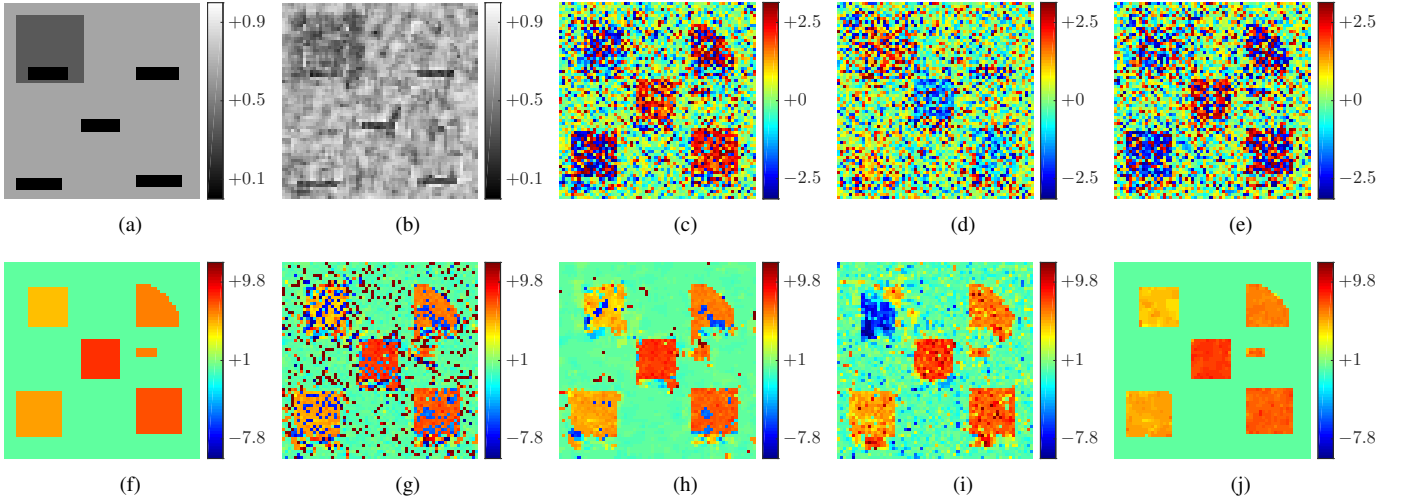


Fig. 1. (a) Starting coherence map, (b) empirical coherence map, (c) first interferogram, (d) second interferogram, (e) third interferogram, (f) original height profile, (g) estimated solution using MLNL approach, (h) estimated solution using MAPNL approach, (i) estimated solution using MCPU approach, (j) estimated solution using the proposed PARISAR approach.

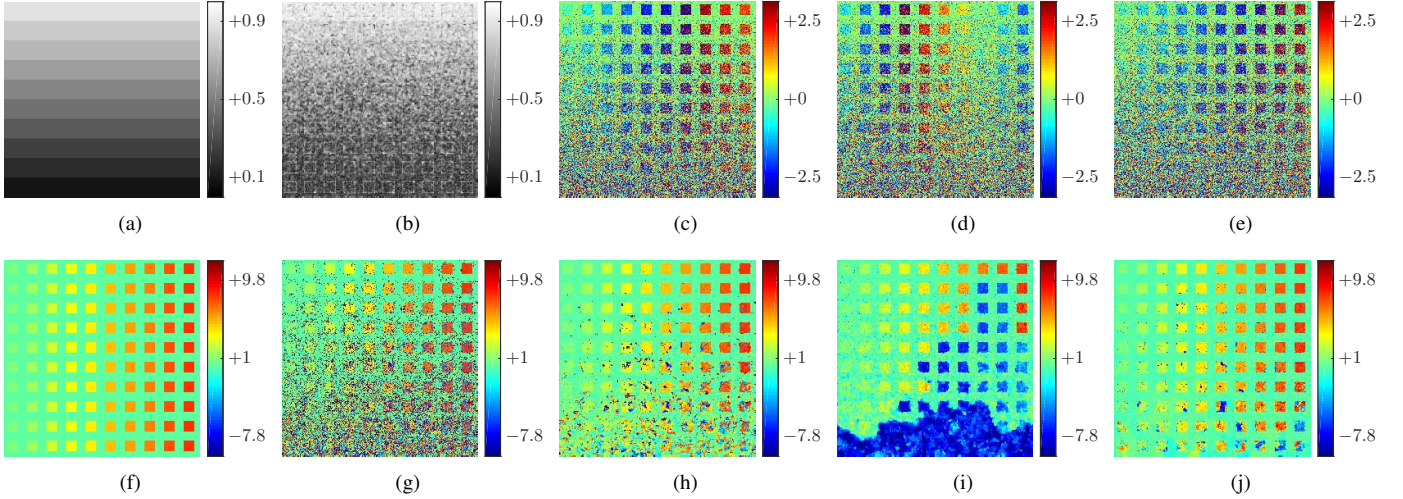


Fig. 2. (a) Starting coherence map, (b) empirical coherence map, (c) first interferogram, (d) second interferogram, (e) third interferogram, (f) original height profile, (g) estimated solution using MLNL approach, (h) estimated solution using MAPNL approach, (i) estimated solution using MCPU approach, (j) estimated solution using the proposed PARISAR approach.

The results obtained using MAPNL, PUMA and PARISAR are shown in Figures 3(f), 3(g) and 3(h), respectively. The first and the last are tested using the whole dataset, PUMA is tested using only the smallest baseline interferogram. It is evident that, even if not ambiguous, a single-channel algorithm as PUMA is penalized by using a single interferogram and therefore fails to correctly retrieve the height, due to the aliasing of fringes. This problem is solved by PARISAR by using all the available channels. Note that exploiting the whole dataset may not be sufficient for correctly retrieving the profile: the difference between results obtained by MAPCorrNL and PARISAR show that the regularization role is important. The quantitative analysis reported in in Tables II and III (see the corresponding lines for the Ghiglia dataset) confirms the visual inspection.

IV. APPLICATION TO SATELLITE SAR IMAGES

To qualitatively evaluate algorithm PARISAR on real satellite SAR images, we considered two datasets: an urban test site (Napoli) and a natural test site (Serre-Ponçon). The two scenes have been acquired using two different sensors, COSMO-SkyMed and ERS, to test the capabilities of the method to work with different radar frequencies (X-band and C-band) and sensors. The systems parameters are summarized in table IV.

A. Urban area: Napoli test case

The first dataset is composed of three 250×250 pixel COSMO-SkyMed Stripmap images acquired close to Naples (Italy) train station. One of the three available interferograms is shown in the first row of Figure 4 together with the mean amplitude (in log scale) and the mean coherence map. The

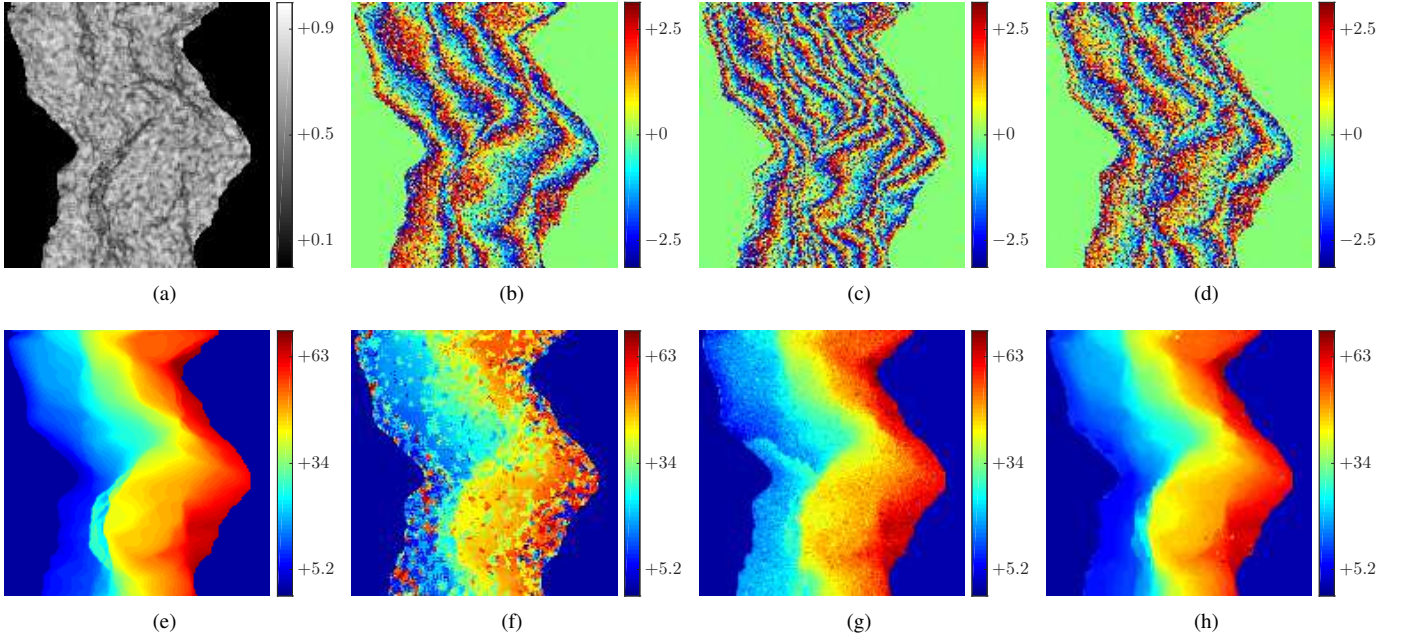


Fig. 3. (a) Empirical coherence map, (b) first interferogram, (c) second interferogram, (d) third interferogram, (e) original height profile, (f) estimated solution using MAPNL approach, (g) estimated solution using PUMA approach, (h) estimated solution using the proposed PARISAR approach.

TABLE IV
INTERFEROMETRIC CONFIGURATIONS OF THE SATELLITE IMAGES

| Dataset | Sensor | R_0 | λ | θ | B_{\perp} |
|----------|--------|---------|-----------|----------|--------------|
| Naples | CSK | 755.190 | 0.03 | 0.62 | [0 517 251]m |
| Serre-P. | ERS2 | 825.669 | 0.05 | 0.40 | [0 36 96]m |

scene is very complex: different structures, with different heights, shapes and reflectivities are present. The phase unwrapping results are shown in the second row. Independent estimation of the height at each pixel leads to a very noisy result (i.e., strong variance of the heights). This is evident from the results of MLNL of Figure 4(e). Considering MAPNL and PARISAR, the regularization reduces these fluctuations without noticeable resolution loss. This reduction is more evident in case of PARISAR (Figure 4(g)): the building structures are retrieved, both in terms of shapes and heights. It is interesting to note the capability of PARISAR in retrieving the low-height circular structures on the left of the scene, while strongly reducing noise: these structures are almost invisible both in the interferogram and in the coherence image (very noisy area). A qualitative evaluation of the reconstruction results can be performed based on the optical images (2D and 3D) of the considered scene provided by Google (Figures 4(d) and 4(h)) and taken at the same time period: from the radar-optical comparison it appears that the structures are correctly retrieved, both in terms of shapes and of relative building heights.

B. Mountainous area: Serre-Ponçon test case

The last 250×250 pixel dataset corresponds to a mountainous area acquired by ERS sensor over Serre-Ponçon (France).

One of the three available interferograms is shown in the first row of Figure 5 together with the mean amplitude (in log scale) and the mean coherence map. This area is challenging due to the presence of very low coherence areas and not regular phase fringes. Phase unwrapping results are shown in the second row. Both MLNL and MAPNL fail at correctly unwrapping the profile. The latter provides a more reliable result although there are several areas that are not correctly unwrapped. Using PARISAR, it is possible to largely improve the results. Wrapping problems are solved and noise is better suppressed.

V. CONCLUSION

A new methodology to improve multi-baseline phase unwrapping has been proposed. Starting from the complete statistical distribution of the interferometric data, the joint exploitation of patch-based approaches and TV regularization for elevation estimation has been discussed. The developed algorithm, named PARISAR, implements a maximum a posteriori estimator with a properly modified likelihood term, by means of a two steps strategy: the first step consists of estimating a covariance matrix at each pixel from the multi-channel images available using a non-local filtering method like NL-SAR; the second step introduces a TV penalty for edge-preserving regularization. PARISAR has been tested on several datasets and compared to other multi-channel algorithms. The quantitative and qualitative analysis has been carried out on three different simulated datasets, in order to validate the effectiveness of the approach in different configurations (various image structures and coherences). A qualitative evaluation has been performed on two satellite image datasets from two different sensors working at different radar frequencies, displaying different spatial resolutions, on a urban and a

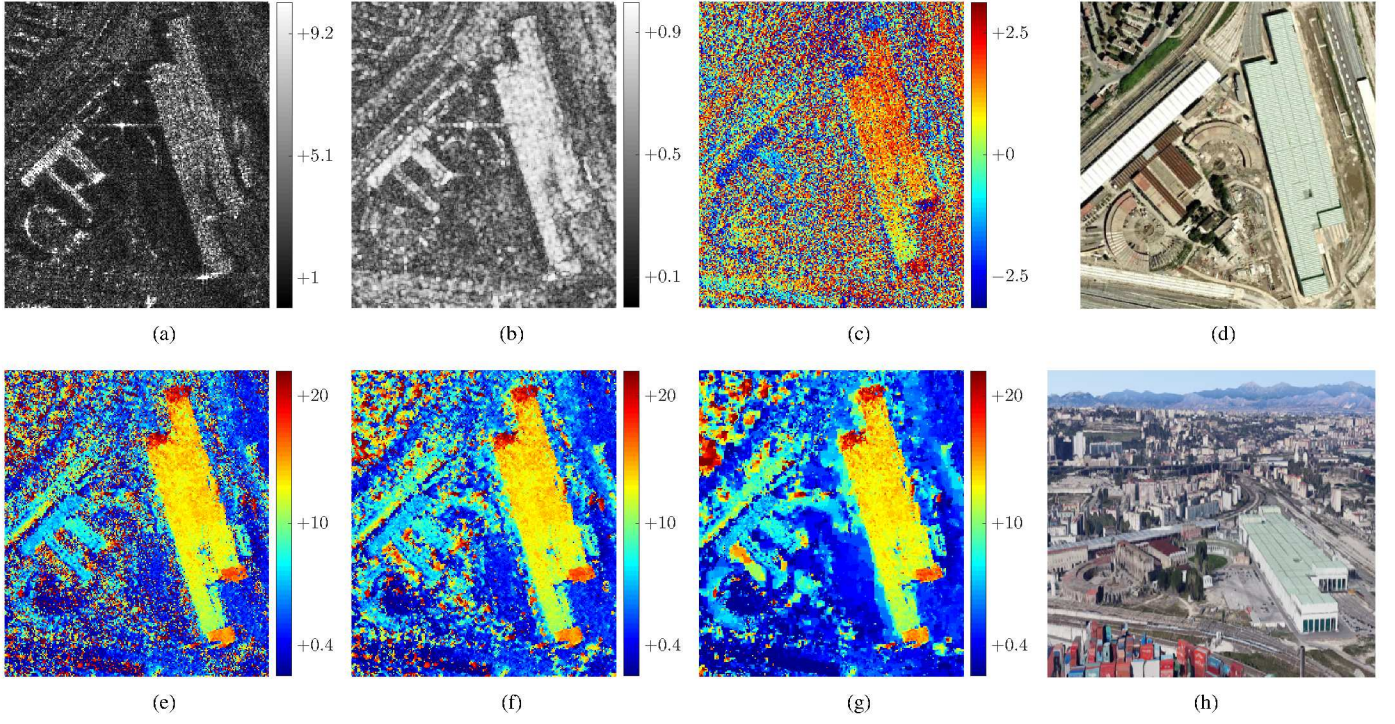


Fig. 4. (a) Mean amplitude image, (b) empirical coherence map, (c) one of the available interferograms, (d) optical image of the considered scene provided by Google, (e) estimated solution using MLNL approach, (f) estimated solution using MAPNL approach, (g) estimated solution using the proposed method PARISAR, (h) 3D optical image of the considered scene provided by Google Earth.

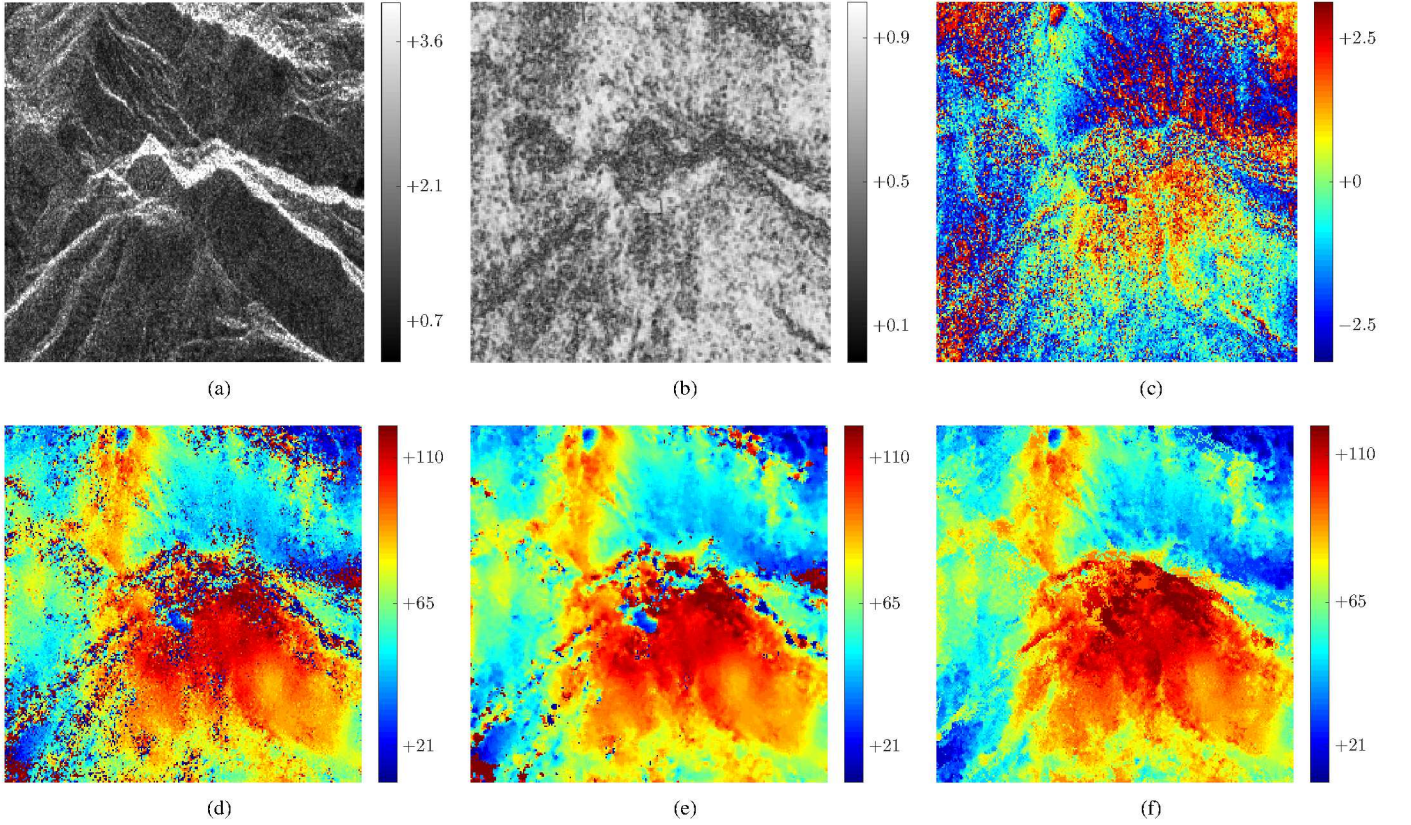


Fig. 5. (a) One of the available amplitude, (b) empirical coherence map, (c) one of the available interferogram, (d) estimated solution using MLNL approach, (e) estimated solution using MAPNL approach, (f) estimated solution using the proposed method PARISAR.

mountainous area. The results in both cases are promising. PARISAR provides sensible elevation profiles, seemingly outperforming other methods. Structural details are preserved while most of the noise is suppressed. At the present stage, the method is able to handle shadows but it does not account for the layover phenomenon: such distortions can be addressed only using a tomographic approach. The investigation of a tomographic-based approach with the proposed framework is the subject of future research.

REFERENCES

- [1] G. Fornaro and V. Pascazio, "SAR Interferometry and Tomography: Theory and Applications," *Academic Press Library in Signal Processing*, Elsevier, 2013.
- [2] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Problem*, vol. 14, pp. R1–R54, August 1998.
- [3] R. Chen, W. Yu, R. Wang, G. Liu, and Y. Shao, "Integrated Denoising and Unwrapping of InSAR Phase Based on Markov Random Fields," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 8, pp. 4473–4485, Aug 2013.
- [4] B. Osmanoğlu, F. Sunar, S. Wdowinski, and E. Cabral-Cano, "Time series analysis of InSAR data: Methods and trends," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 90 – 102, 2016.
- [5] K. Itoh, "Analysis of the phase unwrapping problem," *Appl. Opt.*, vol. 21, no. 14, 1982.
- [6] V. Pascazio and G. Schirinzì, "Multifrequency InSAR height reconstruction through maximum likelihood estimation of local planes parameters," *IEEE Transactions on Image Processing*, vol. 11, no. 12, pp. 1478–1489, 2002.
- [7] G. Ferraiuolo, F. Meglio, V. Pascazio, and G. Schirinzì, "DEM Reconstruction Accuracy in Multichannel SAR Interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 191–201, 2009.
- [8] F. Li and R. Goldstein, "Studies of multibaseline spaceborne interferometric synthetic aperture radars," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 1, pp. 88–97, 1990.
- [9] A. Ferretti, C. Prati, and F. Rocca, "Multibaseline phase unwrapping for InSAR topography estimation," *Nuovo cimento della Società italiana di fisica. C, Geophysics and space physics*, vol. 24, no. 1, pp. 159–176, 2001.
- [10] F. Gini and F. Lombardini, "Multibaseline cross-track SAR interferometry: a signal processing perspective," *IEEE Aerospace and Electronic Systems Magazine*, vol. 20, no. 8, pp. 71–93, Aug 2005.
- [11] G. Fornaro, A. M. Guarnieri, A. Paucillo, and F. De-Zan, "Maximum likelihood multi-baseline SAR interferometry," *IEEE Proceedings - Radar, Sonar and Navigation*, vol. 153, no. 3, pp. 279–288, June 2006.
- [12] H. Liu, M. Xing, and Z. Bao, "A Cluster-Analysis-Based Noise-Robust Phase-Unwrapping Algorithm for Multibaseline Interferograms," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 494–504, 2015.
- [13] H. Liu, M. Xing, and Z. Bao, "A Novel Mixed-Norm Multibaseline Phase-Unwrapping Algorithm Based on Linear Programming," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 5, pp. 1086–1090, 2015.
- [14] D. Chirico and G. Schirinzì, "Multichannel interferometric sar phase unwrapping using extended kalman smoother," *International Journal of Microwave and Wireless Technologies*, vol. 5, pp. 429–436, 6 2013.
- [15] M. Schmitt and U. Stilla, "Maximum-likelihood estimation for multi-aspect multi-baseline SAR interferometry of urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, no. 0, pp. 68 – 77, 2014.
- [16] A. Shabou, F. Baselice, and G. Ferraioli, "Urban digital elevation model reconstruction using very high resolution multichannel InSAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4748–4758, 2012.
- [17] 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," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 24–36, 2005.
- [18] G. Ferraiuolo, V. Pascazio, and G. Schirinzì, "Maximum a posteriori estimation of height profiles in InSAR imaging," *IEEE Geoscience and Remote Sensing Letters*, vol. 1, no. 2, pp. 66–70, 2004.
- [19] G. Nico, G. Palubinskas, and M. Datcu, "Bayesian approaches to phase unwrapping: theoretical study," *IEEE Transactions on Signal Processing*, vol. 48, no. 9, pp. 2545–2556, 2000.
- [20] F. Baselice, G. Ferraioli, V. Pascazio, and G. Schirinzì, "Contextual information-based multichannel synthetic aperture radar interferometry: Addressing DEM reconstruction using contextual information," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 59–68, July 2014.
- [21] G. Ferraioli, A. Shabou, F. Tupin, and V. Pascazio, "Multichannel phase unwrapping with graph cuts," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 3, pp. 562–566, 2009.
- [22] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger, "NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In) SAR denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2021–2038, 2015.
- [23] G. Gilboa, N. Sochen, and Y. Zeevi, "Variational denoising of partly textured images by spatially varying constraints," *IEEE Transactions on Image processing*, vol. 15, no. 8, pp. 2281–2289, 2006.
- [24] C. Soutour, C.-A. Deledalle, and J.-F. Aujol, "Adaptive Regularization of the NL-Means: Application to Image and Video Denoising," *Image Processing, IEEE Transactions on*, vol. 23, no. 8, pp. 3506–3521, Aug 2014.
- [25] J. W. Goodman, "Some fundamental properties of speckle," *JOSA*, vol. 66, no. 11, pp. 1145–1150, 1976.
- [26] J. Fan, M. Farnen, and I. Gijbels, "Local maximum likelihood estimation and inference," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 60, no. 3, pp. 591–608, 1998.
- [27] J. Polzehl and V. Spokoiny, "Propagation-separation approach for local likelihood estimation," *Probability Theory and Related Fields*, vol. 135, no. 3, pp. 335–362, 2006.
- [28] C.-A. Deledalle, L. Denis, and F. Tupin, "Iterative weighted maximum likelihood denoising with probabilistic patch-based weights," *IEEE Transactions on Image Processing*, vol. 18, no. 12, pp. 2661–2672, Dec 2009.
- [29] J.-S. Lee, "Digital image smoothing and the sigma filter," *Computer vision, graphics, and image processing*, vol. 24, no. 2, pp. 255–269, 1983.
- [30] J.-S. Lee, "Refined filtering of image noise using local statistics," *Computer graphics and image processing*, vol. 15, no. 4, pp. 380–389, 1981.
- [31] C.-A. Deledalle, L. Denis, G. Poggi, F. Tupin, and L. Verdoliva, "Exploiting patch similarity for SAR image processing: the nonlocal paradigm," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 69–78, 2014.
- [32] F. Baselice, A. Budillon, G. Ferraioli, V. Pascazio, and G. Schirinzì, "Multibaseline SAR Interferometry from Complex Data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 7, pp. 2911–2918, July 2014.
- [33] H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1333–1336, 2003.
- [34] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, Wiley, 1998.
- [35] J. Bioucas-Dias and G. Valadao, "Phase unwrapping via graph cuts," *Image Processing, IEEE Transactions on*, vol. 16, no. 3, pp. 698–709, 2007.