# Bayesian Approaches to Phase Unwrapping: Theoretical Study

Giovanni Nico, *Member, IEEE*, Gintautas Palubinskas, and Mihai Datcu

***Abstract*—The problem of phase unwrapping of two-dimensional (2-D) phase signals has gained a considerable interest in recent years. It deals with the problem of estimating (reconstructing) an absolute phase from the observation of its noisy principal (wrapped) values. This is an ill-posed problem since many possible solutions correspond to a given observation. Many phase unwrapping algorithms have been proposed relying on different constraints for the phase signal sampling process or the nature (e.g., smoothness, regularity) of the phase signal. We look at these algorithms from the Bayesian point of view (estimation theory) and analyze the role of the *prior assumptions*, studying their equivalencies to the regularization constraints already used. This study lead to the development of the two new phase unwrapping algorithms, presented in the last section of this paper, which are able to work in quite difficult conditions of aliasing and noise. The theoretical study of the analyzed schemes is illustrated by some experiments on synthetic phase signals.**

***Index Terms*—Bayesian modeling, inference, interferometric phase, phase unwrapping.**

## I. Introduction

A VARIETY of imaging systems use coherent radiation to illuminate objects. The scattered field returned back to the sensor carries information about the physical and geometrical properties of the imaged objects. For example, in optical interferometry, the absolute phase reconstructed from the interferometric pattern contains information on shape, displacements, and vibrations of the object's surface [1], [2]. Other examples are in adaptive optics [3]–[6], magnetic resonance imaging [7], diffraction tomography [8], [9], synthetic aperture radar (SAR) interferometry [10]–[13], homomorphic signal processing [14], and solid-state physics [15].

In all previously enumerated cases, the information is carried by the phase of the sensed signals. The phase can be observed only in the principal interval $(-\pi, \pi]$, and it is said to be "wrapped." To extract the interesting information, e.g., height surface (digital elevation model—DEM), the absolute phase should be estimated from the wrapped observations. This is the task of phase unwrapping (PU). An overview of the field is given by Ghiglia and Pritt in [16]. In the general case, it is not possible to unambiguously recover the unknown absolute phase from the measured wrapped phase, which is usually corrupted by the noise [17]–[19]. Consequently, additional information is required to constrain the PU problem, e.g., we may know that the surface to be reconstructed fullfils some smoothness or regularity assumptions. Early proposed PU algorithms use local analysis of the wrapped phase signal differences (gradients) or apply edge detection techniques to recover the fringes, $2\pi$ phase jumps [20], [21]. Due to the presence of noise and the signal undersampling (aliasing), these methods are limited in the accuracy of the provided solutions. A well-known approach to solve the inconsistency of the estimated phase gradients is the insertion of branch cuts between residues of opposite polarity and then preventing the integration path from crossing them. These algorithms are called path-following algorithms. Another group of developed algorithms are path independent. They are based on a least-squares estimate of the phase by minimizing the squared distance between an estimate of the absolute phase gradient and the true gradient of the unknown absolute phase. They are minimum-norm algorithms. It was demonstrated that this method is equivalent to the solution of Poisson's equation with Neumann boundary conditions and can be obtained using the Green's function. Based on the methods presented above, a number of PU algorithms have been implemented [22]–[32]. Only a few researchers approached the problem of PU from the point of view of estimation theory [33]–[36].

We analyze several of these methods in which the hypotheses (assumptions) for the absolute phase, expressed in the form of *a priori* surface models, are exploited. The analysis is done in the context of Bayesian inference, thus enabling a systematic investigation of different algorithms for PU: identification of their common properties and reformulation of them as "model based" solutions. This theoretical analysis is followed by illustrative experiments on synthetic phase signals. We use Gaussian phase surfaces because they are used usually as a means of evaluating PU algorithms under controlled conditions (see, for example, [37]). The PU algorithm robustness is tested on phase signals with and without aliasing, corrupted with a various amounts of noise (see Fig. 1).

This paper has the following structure. In Section II, we introduce the PU problem. We analyze its ill-posed nature and the two-fold signal phenomena (aliasing and noise), which hinder obtaining the solution of PU using simple deterministic methods. In Section III, we present the basics of Bayesian inference, maximum entropy principles, and their application in model-based parameter estimation. Further, we overview the basic principles of regularization used for surface reconstruction. In Section IV, we present and analyze the most important PU algorithms from the Bayesian perspective. We exemplify

Manuscript received June 7, 1999; revised May 12, 2000. The associate editor coordinating the review of this paper and approving it for publication was Prof. Colin F. N. Cowan.

G. Nico is with the Space Applications Institute, Joint Research Centre TP.272 of the European Commission, Ispra, Italy (e-mail: Giovanni.Nico@jrc.it).

G. Palubinskas and M. Datcu are with the German Aerospace Center (DLR), Remote Sensing Technology Institute, Oberpfaffenhofen, Wessling, Germany (e-mail: Gintautas.Palubinskas@dlr.de; Mihai.Datcu@dlr.de).

Publisher Item Identifier S 1053-587X(00)06668-X.

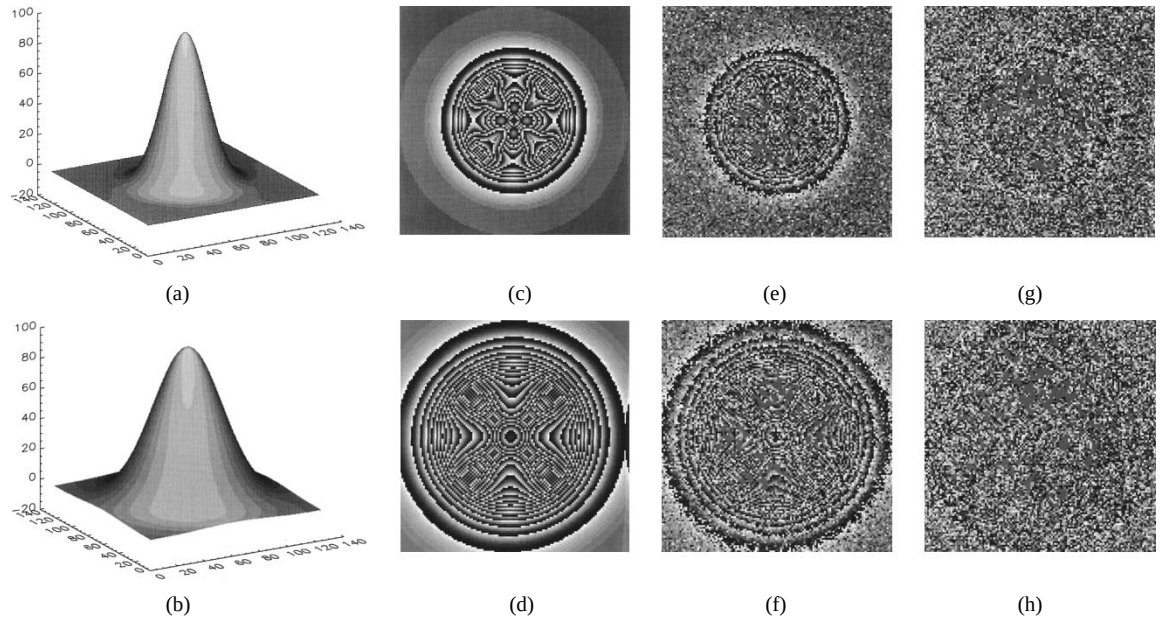


Fig. 1. Gaussian absolute phase surface with aliasing occurring where adjacent phase samples differ more than $\pi$ is depicted in (a). Interferograms are, respectively, (c) noiseless, (e) with an interferometric noise corresponding to a coherence $\gamma = 0.9$, (g) noise with $\gamma = 0.6$. Gaussian absolute phase surface without aliasing is shown in (b). All adjacent phase samples differ less than $\pi$. Interferograms are, respectively, (d) noiseless, (f) noise with $\gamma = 0.9$, (h) noise with $\gamma = 0.6$.

the role and power of the *a priori* models used for absolute phase reconstruction in adverse observation conditions, both with aliasing and noise. The article is concluded in Section V with a presentation of the advantages of using Bayesian modeling for the formulation of the PU problem.

## II. Phase Unwrapping: Notation and Definitions

The interferogram is a complex signal $\exp\{\imath f\}$ (e.g., in interferometric SAR—InSAR), where $f$ is the absolute phase and carries the information about the observed process and $\imath = \sqrt{-1}$. The absolute interferometric phase $f$ has to be determined from noisy samples of the interferogram defined on a regular grid $I$

$$\exp\{\imath f(i,j)\} = \exp\{\imath f(x_0 + i\Delta x, y_0 + j\Delta y)\} \tag{1}$$

where $x_0, y_0$ are the spatial coordinates, and $\Delta x, \Delta y$ are spacing steps, respectively. The integers $(i,j) \in I$, with $i = 1,\ldots,N_x, j = 1,\ldots,N_y$, indicate the phase samples in the two dimensions. The wrapping operator $\mathcal{W}$ is defined as

$$\begin{aligned}\mathcal{W} &: \exp\{\imath f(i,j)\} \in \mathbb{C} \rightarrow g(i,j) \\ &= \arctan\left(\frac{\Im[\exp\{\imath f(i,j)\}]}{\Re[\exp\{\imath f(i,j)\}]}\right) \in (-\pi, \pi].\end{aligned} \tag{2}$$

It extracts the observed wrapped phase $\{g(i,j)\}$ from the complex interferogram $\exp\{\imath f(i,j)\}$.

The problem of PU consists in estimating the absolute phase $\{f(i,j)\}$ from the observed noisy wrapped phase $\{g(i,j)\}$

$$g(i,j) = \mathcal{W}[f(i,j) + n(i,j)] \tag{3}$$

where $\{n(i,j)\}$ is the field describing the noise in the interferogram [17], [18]. This is an ill-posed problem, meaning that many solutions may correspond to a given interferogram. PU is equivalent to the classical ill-posed problem of reconstructing a vector from its projection (see, for example, [23]). Assumptions on the underlying sampling process or on the nature of the desired solution are needed to tackle the PU. For example, Itoh [38] demonstrated that when a one-dimensional (1-D) phase signal is sampled in such a way that its gradient $\nabla f = f(i+1) - f(i)$ satisfies the condition (no aliasing)

$$|\nabla f| < \pi \tag{4}$$

the following relation

$$\nabla f = \mathcal{W}[\nabla g] \tag{5}$$

is verified. We will call (4) the correct sampling condition. Relation (5) enables us to unwrap a phase signal by simply integrating the quantities $\nabla f$. Thus, most of the algorithms for phase unwrapping are based on a two-step procedure: i) the estimation of the unwrapped phase gradient $\nabla\hat{f}$ from the wrapped observations $\mathcal{W}[\nabla g]$ and ii) computation of the absolute phase $\hat{f}$ by integration of the previous estimate $\nabla\hat{f}$. This approach always produces a result but not always a satisfactory one. However, in the 1-D case, there is no possibility to use the observed phase $g$ to verify whether the condition (4) is satisfied or not.

A direct extension of the Itoh's algorithm to two-dimensional (2-D) phase arrays gives unacceptable results. Itoh's algorithm fails because of the undersampling of the phase signal (aliasing) and due to noise, where condition (4) is not always satisfied. This gives rise to residues. A residue is an inconsistency of the estimated phase gradient $\nabla\hat{f}$ and occurs where the sum of the following quantities

$$\nabla\hat{f} = \mathcal{W}\left[\nabla_{x(y)} g(i,j)\right] \tag{6}$$

calculated along the smallest closed path (a square of $2 \times 2$ pixels) is different from zero ($x$ or $y$ stands for direction in the coordinate system). Residues are a property of 2-D PU and have

no analogy in one dimension. In the absence of corrupted regions [(4) is verified], the unwrapped phase gradient $\nabla \hat{f}$ is irrotational [39]; thus, the absolute phase can be recovered using a path independent integral. If the unwrapped phase has high variations and it is not correctly sampled (it is aliased), or the signal-to-noise ratio is low, residues are present. Thus, the solution will be dependent on the integral path and will no longer be unique. The method to avoid the path dependent solutions is based on insertion of cut lines between residues of opposite polarity and not allowing the integration path to intersect these cuts. In this way, the path dependence of PU results, characterizing Itoh's algorithm, is avoided by selecting an appropriate integration path not crossing branch cuts. Recently, an algorithm has been proposed that guides the unwrapping along reliable paths [40]. Reliability is measured in terms of noise and unwrapping consistency. We will again come to this topic in Section IV-C.

All PU algorithms analyzed in this paper estimate the gradients of the absolute phase in different ways. The absolute phase is found then from these gradients by integration.

For the testing of PU algorithms, we use Gaussian phase surfaces (see Fig. 1) because they are well suited as a means of evaluating PU algorithms under controlled conditions [37]. The phase slope, i.e., the difference between adjacent pixels in the phase array, depends on the sample spacing. In particular, we will use both Gaussian phase surfaces whose absolute slope is less than $\pi$ and surfaces with an absolute phase slope greater than $\pi$. In the first case, the phase signal has been sampled following the Nyquist theorem, whereas the second phase signal is said to be aliased, i.e., it is undersampled. We will also consider the noisy versions of these phase signals. The amount of noise in interferograms depends on the interferometric coherence that controls the quality infomation of a phase signal [17], [18]. Therefore, the PU algorithm's robustness is tested on phase signals with and without aliasing and corrupted with a various amounts of noise.

## III. Bayesian Modeling Theory

In this section, we will introduce two methods of inference, the model-based parameter estimation based on Bayesian modeling and, finally, some prior models used for surface regularization.

### A. Methods of Inference

This section gives a short introduction into two methods of inference: Bayesian inference and maximum entropy principle.

The Bayesian approach arises from the fundamental approach of considering the probability theory as a set of normative rules for conducting inference. This approach to probability theory was first outlined clearly by Cox in 1946 [41]. Supposing that degrees of plausibility are to be represented by real numbers on a scale, he found the conditions that such a calculus can be consistent in the sense that if two different methods of calculation are permitted by the rules, then they should yield the same result. These consistency requirements took the form of two functional equations whose solutions uniquely determined the fundamental probability calculus rules (the basic product and sum rules of the probability theory) from which all the results of the probability theory can be derived. Thus, the fundamental equations of probability theory are uniquely determined as rules of inference. These results are based only on some very elementary requests on consistency that make no reference to random experiments; therefore, it is legitimate to assign probabilities to any clearly stated proposition. Thus, the conventional interpretation of probability as frequency in a random experiment is included as a special case of probabilistic inference for certain kinds of propositions or prior information that make explicit statements about the outcomes of an experiment, whatever it is. Moreover, the "frequency" approach could never lead to the "inferential" approach to probability simply because the rationales behind the first are merely deductive.

If the probability theory is a set of normative rules for conducting inference, then the Bayes' theorem

$$p(A \mid B) = \frac{p(A)p(B \mid A)}{p(B)} \tag{7}$$

is to be seen as a method of translating information into a probability assignment, where $A$ and $B$ are some propositions and $p(B) > 0$. Therefore, in the Bayesian approach, we can assign to any proposition the probability density function, that is, to model it.

Another method that accomplishes such task of modeling is the principle of maximum entropy (ME). ME is a general method of inference about unknown probability densities subject to a set of constraints. In terms of the inference of the probability densities, suppose one has a system that has a set of states $s = \{s_i\}$ having the unknown probabilities $q(s_i)$. Additional information about the probabilities of the states is also available in term of constraints on the distribution $q(\cdot)$: values of certain expectations, for example. These expectations constrain the space of the solutions to a certain domain (preferable convex), but usually, there remains an infinite set of distributions that are not ruled out by the constraints. The principle of ME states that of all distributions $q(\cdot)$ that satisfy the constraints, one should choose the one with the largest entropy

$$\mathcal{H}(q) = -\sum_i q(s_i)\log(q(s_i)). \tag{8}$$

The ME solution will be understood as the most uniform solution. However, ME fails to take noise into account: a factor with which classical methods deal. For the case when we have both noise and prior information, one must use the Bayesian method, incorporating as limit cases both the classical approach and the ME method. Equation (7) shows how the prior probability $p(A)$ is changed into the posterior probability $p(A \mid B)$ as a result of acquiring new information $B$.

The following sections will show how the theory presented above can be used in solving the PU problem.

### B. Model-Based Estimation

The problem of absolute phase reconstruction (estimation), as presented in Section II, uses a deterministic model for the PU solution. We consider a stochastic point of view and approach

it from the Bayesian estimation perspective. The absolute phase $f$ is obtained as a maximum *a posteriori* (MAP) estimate

$$\hat{f} = \arg\max_f \{\mathcal{P}(f \mid g)\} \tag{9}$$

where $\mathcal{P}$ is the sum of two terms: One ($\mathcal{L}$) measures the faithfullness of the estimate to the data, and the second is a regularization term $\mathcal{R}$:

$$\mathcal{P}(f \mid g) = \mathcal{L}(g \mid f) + \lambda \mathcal{R}(f) \tag{10}$$

where $\lambda$ is a hyperparameter weighting the two terms.

The Bayesian approach allows us to give the probabilistic interpretation for the terms $\mathcal{L}$ and $\mathcal{R}$. Without loss of generality, we can write

$$\mathcal{L}(g \mid f) = -\log p(g \mid f) \tag{11}$$

where $p(g \mid f)$ is the probabilistic model of the uncertainty of our observation process: the likelihood function. The likelihood function also includes the deterministic model, which describes how the observations have been obtained; see (3). Therefore, the likelihood describes the wrapped phase $g$, which for InSAR has a hypergeometrical distribution [39] and is difficult to handle analytically. The usual way is that the likelihood of the data is enforced by using a quadratic functional of the difference between the reconstructed $\nabla f$ and the measured $\mathcal{W}[\nabla g]$, because of (4) and (5). A $\chi^2$ measure of the error between the reconstructed and the observed data is assumed (e.g., see Section IV-A). This is, of course, an approximation assuming an additive Gaussian white distribution for the noise in the data. The regularization term represents the restriction we impose, using some prior knowledge, on the solution

$$\mathcal{R}(f) = -\log p(f) \tag{12}$$

where $\mathcal{R}(f)$ is the *a priori* model. The Bayesian frame allows to "compose" the likelihood and the *a priori* information and infer the *a posteriori* model

$$p(f \mid g) = \frac{p(g \mid f)p(f)}{p(g)} \tag{13}$$

which is linked to the operator $\mathcal{P}$

$$\mathcal{P}(f \mid g) = -\log p(f \mid g). \tag{14}$$

Several remarks are important. Due to the nature of the wrapping operator, which describes the interferogram formation process, the likelihood is a periodic function [35]. In our discussion, we will consider the likelihood to be well approximated by a Gaussian function around the mode. The inference of the posterior distribution from the likelihood and the *a priori* models used is qualitatively presented in Fig. 2.

The *a priori* information, i.e., the regularization term, is defined as a sum of potentials (relation to the Gibbs distribution), which are, in general, functions of a derivative of the absolute phase. Exception are made by maximum entropy and fractal models.

### C. Prior Models

The phase unwrapping is an ill-posed problem. In order to obtain a unique solution, it must be regularized, assuming a prior model for the unknown surface. The prior model is selected either based on prior knowledge of the imaged scene or asserting certain desired smoothness for the solution. In the Bayesian frame, the model selection is a special case of inference; however in this paper, we do not deal with this aspect.

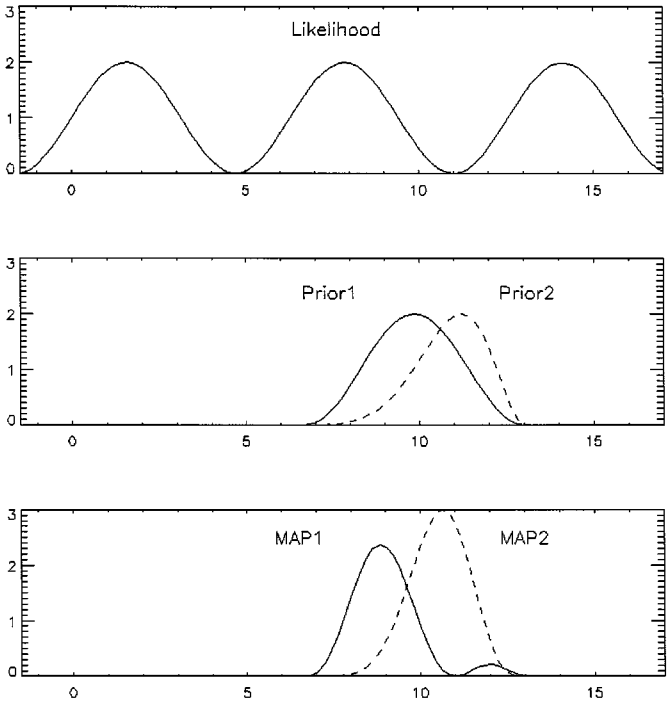



Fig. 2. Inference of posterior distribution (lower diagram) from the likelihood function (upper) and two prior models (middle).

A simple regularization model (first order) is the potential energy constrained in a stretched membrane; it penalizes the variations from $\nabla_x f = 0$ and $\nabla_y f = 0$ by minimizing

$$\mathcal{R}(f) = \sum\sum[(\nabla_x f)^2 + (\nabla_y f)^2] \tag{15}$$

where the sum is computed in the neighborhood of the current spatial coordinate.

Another common regularization model (second order) constraints the surface to be an isotropic thin plate. The regularization is specified in terms of the second derivatives:

$$\mathcal{R}(f) = \sum\sum[(\nabla_x \nabla_x f)^2 + (\nabla_x \nabla_y f)^2 + (\nabla_y \nabla_x f)^2 + (\nabla_y \nabla_y f)^2]. \tag{16}$$

The previous two regularization models play an important role in different vision problems, e.g., surface reconstruction problems and shape from shade (see [42] and reference therein). They are of central importance in interferometry because the measurement are directly functions of surface gradients.

In general, the estimates of $\nabla_x f$ and $\nabla_y f$ do not correspond to the true gradients of the surface $f$. This is the surface consistency problem resulting from aliasing and poor signal-to-noise ratio. The estimated gradients (6) must satisfy the surface consistency condition

$$\oint_f (\nabla_x f\, dx + \nabla_y f\, dy) = 0 \tag{17}$$

as already discussed in Section II. This leads to another type of regularization model for absolute phase by means of the functional

$$\mathcal{R}(f) = \sum_{(i,j)} [\nabla_x f(i,j) + \nabla_y f(i+1,j) - \nabla_x f(i,j+1) - \nabla_y f(i,j)]^2 \quad (18)$$

which, with the value equal to zero, constrains the phase gradients to be consistent. This model includes the observation inside because of (6). In the following, we will refer to (18) as Bayesian estimation through residue analysis, i.e., the prior is a probability density function (pdf) with two values: one for the rotational case and a small value close to zero for the consistency case. Therefore, the Bayesian framework allows us to treat both types of prior information: regularization terms and consistency contraint in the same manner in the functional (10).

The principle of maximum entropy provides one more regularization model. This is a nonlinear approach (see Section III-A), which extracts the information directly from the data. The regularization term is

$$\mathcal{R}(f) = -\sum \bar{f} \log \bar{f} \quad (19)$$

where $\bar{f} = f / \sum f$.

Many natural surfaces have been demonstrated to be well approximated by fractal models. The most popular model used in this respect is the fractional Brownian motion (fBm) model [57]. This model describes a signal $B(t)$ characterized by the fact that its increments between two moments of time $t$ and $t+s$ have a variance proportional to a power $2H$ of the time lag $|s|$

$$\mathcal{E}\{(B_H(t+s) - B_H(t))^2\} \approx |s|^{2H} \quad (20)$$

where $\mathcal{E}$ stands for the expectation of the random variable, and the parameter $H$ is called the *Hurst exponent* $(0 < H < 1)$ and is related to the fractal dimension $D$ of the corresponding fractal by $D = n+1-H$, where $n$ is the topological dimension of the space in which the fractal object is represented ($n = 2$ for fBm surfaces). Note that fBm signals are nonstationary. This makes their analysis both from the theoretical and from the practical point of view quite difficult. Closed form does not exist. However, the increments of fBm process are strict sense stationary, and the random variable $(B_H(t+s) - B_H(t))/|s|^H$ is Gaussian and has zero mean, which allows us to use it as a prior model in Bayesian estimation. Thus, the fractal models can be seen as a general class of regularization that includes, as a partial case, the first-order regularization; see (15).

In the following section, we will try to show what models or assumptions were used in deriving some of the PU algorithms.

## IV. Bayesian Approaches to PU

### A. Least-Squares Estimation

The most serious drawback of Itoh's algorithm is its depedence on the integration path. This problem is solved by the least-squares estimation (LSE) approach to PU. The LSE scheme gives a solution by minimizing the functional $\mathcal{P}_{\text{LSE}} = \mathcal{L}_{\text{LSE}}$ with $\mathcal{R}$ ignored

$$\mathcal{L}_{\text{LSE}} = \sum_{(i,j)\in I_x} (\nabla_x f(i,j) - \mathcal{W}[\nabla_x g(i,j)])^2 + \sum_{(i,j)\in I_y} (\nabla_y f(i,j) - \mathcal{W}[\nabla_y g(i,j)])^2 \quad (21)$$

which depends on the phase derivatives fields $\{\nabla f(i,j)\}$ and $\{\nabla g(i,j)\}$. Prior knowledge is not used in (21). This estimate is biased when obtained in a noisy environment [43]. The LSE functional is a special case of that produced by the Bayesian approach. It corresponds to take only the likelihood term (11). There are many PU algorithms based on the LSE scheme [27]–[32], [44], [45]. Differences are only in the way the LSE functional is minimized, and as a consequence, they differ in efficiency and robustness. Our aim is to examine the *a priori* assumption used to build the functional. The assumption on the sampling process is that the desired unwrapped phase verifies (4) everywhere on the grid $I$. This is the major drawback in the LSE approach because it is equivalent to stating that the interferogram has no residues. In particular, the unwrapping becomes much more complicated when aliasing is present. In this case, the LSE approach fails to give the correct solution even in the absence of noise, as can be observed from Fig. 3. This means that every true phase slope $\nabla f(i,j)$ whose value is outside $(-\pi, \pi]$ cannot be correctly estimated. In Fig. 4(a) and (b), the LSE solution profiles are compared with the original one. The LSE PU algorithm performs well for nonaliased signals with quite low noise. Aliased signals (even noisefree) destroy the LSE solution dramatically.

The LSE solution is not restricted to be obtained by adding $2\pi$-cycles to the observed gradients. The estimated quantities $\nabla f(i,j)$ can assume any value that best fit the wrapping of the wrapped phase gradients $\mathcal{W}[\nabla_y g(i,j)]$. This gives the possibility to compare the "interferogram," which is obtained by wrapping the LSE solution, with the original one. In this way, a direct hint on the goodness of the solution can be obtained. This possibility does not exist in "integer" PU algorithms as Itoh's one. Finally, if one wishes a solution whose interferogram coincides with the original one, the congruence operation can be applied to the LSE solution [46].

To overcome the LSE limits, a functional built using more general assumptions is needed. In certain cases, the LSE scheme can be still valid if a different, more general and reliable, estimate of phase gradients can be obtained (see, for example, [47]–[49]). Another possibility consists of deriving a new functional that is based on less restrictive assumptions, therefore exploiting much more information contained in the interferogram. In the following sections, some of these attempts will be described.

### B. First-Order Regularization

In this section, an algorithm is described that solves PU by using the first-order regularization (15). The algorithm gives an estimate of the derivative $\nabla f(i,j)$. The functional used by Marroquin and coworkers [33] consists of the classical likelihood and regularity terms

$$\mathcal{P} = \mathcal{L}_{\text{LSE}} + \lambda \left\{ \sum_{(i,j)\in I_x} (\nabla_x f(i,j))^2 + \sum_{(i,j)\in I_y} (\nabla_y f(i,j))^2 \right\}. \quad (22)$$

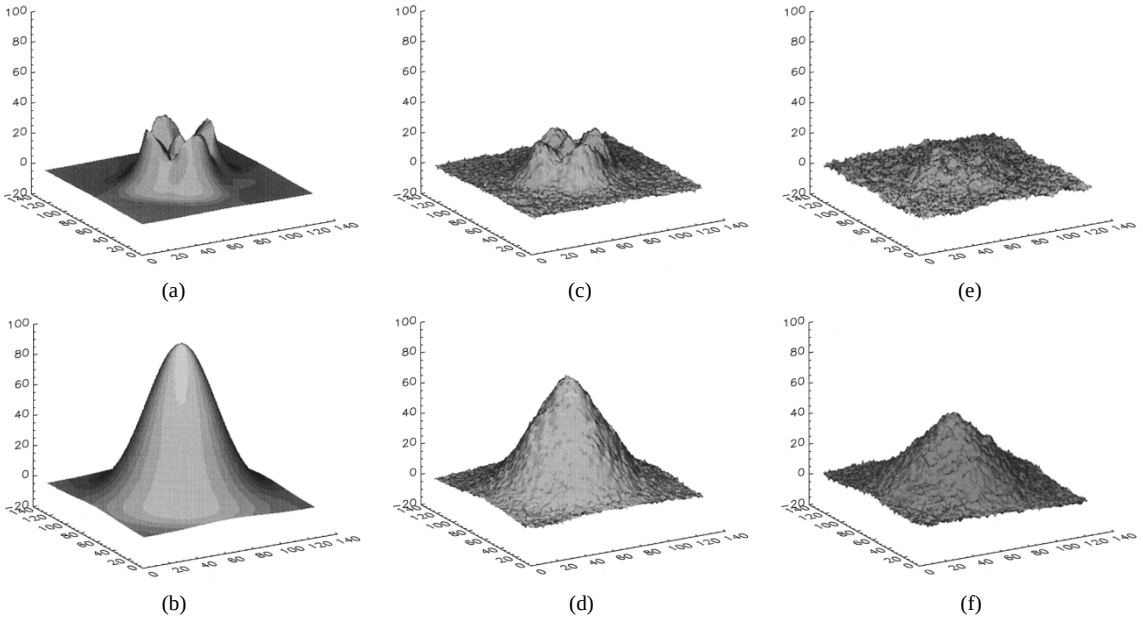


Fig. 3. LSE solutions relative to the the phase signals depicted in Fig. 1 are shown. Aliased phase signal: (a) noiseless, (c) noise with $\gamma = 0.9$, (e) noise with $\gamma = 0.6$. Phase signal without aliasing: (b) noiseless, (d) noise with $\gamma = 0.9$, (f) noise with $\gamma = 0.6$.

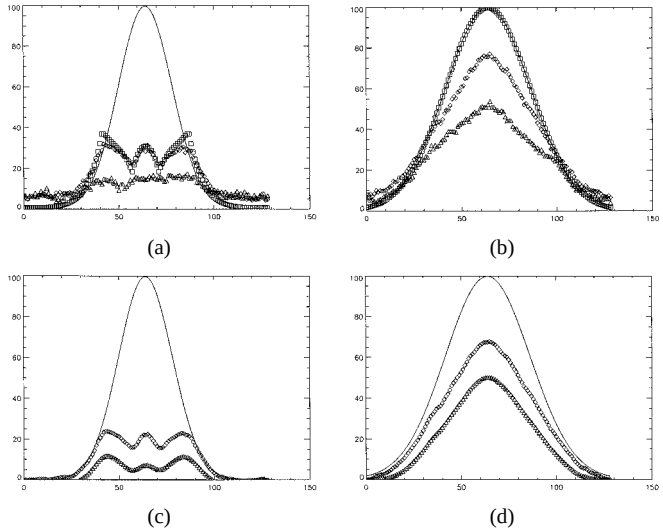


Fig. 4. LSE and first-order regularization solution deformations. In (a), the original aliased phase signal profile, with the corresponding LSE solution profiles, are depicted. The data refer to the noiseless ($\square$), noise with $\gamma = 0.9$ ($\diamond$), and noise with $\gamma = 0.6$ ($\triangle$) cases, respectively. In (b), the corresponding data for the nonaliased phase signal are depicted. The first-order regularizaton solution profiles are shown in (c) for noisy aliased interferogram ($\gamma = 0.9$) and (d) for noisy nonaliased interferogram ($\gamma = 0.9$), respectively. The true phase signal profiles are compared with those obtained by the first-order regularization approach, using different $\lambda$ values: $\lambda = 1(\diamond)$, $\lambda = 0.1(\triangle)$.

Equation (22) constitutes a generalization of the LSE algorithm (21). The two functionals coincide when we do not take into account the second term embedding the *a priori* knowledge.

Even if the functional (22) is able to inject into the solution some *a priori* knowledge about its regularity, it has the same limits as LSE because of the likelihood term. In Fig. 4(c) and (d), the results obtained by applying this algorithm to the interferogram of the Gaussian phase surfaces depicted in Fig. 1(e) and (f) for different chosen $\lambda$-values are reported. As can be seen, this algorithm shares the same drawbacks as the LSE approach because it makes the same fundamental assumption on the phase sampling process. Moreover, the algorithm described in this section filters the interferogram. In fact, the solution becomes less noisy and flatter as $\lambda$ grows. This causes a loss of information as shown in Fig. 4(c) and (d). Practically, the performace of this algorithm is very similar (in tested conditions) to that of LSE.

### *C. Bayesian Estimation through Residue Analysis: First-Order Regularization*

Another way to generalize the LSE approach within the Bayesian regularization scheme is described in this section. The likelihood term requires the absolute phase gradients $\nabla_x f$ and $\nabla_y f$ match the observed wrapped differences, using the same LSE functional, whereas the regularization term constrains these gradients to be consistent by means of the functional (18). The complete proposed regularization functional is [50] $\mathcal{P} = \mathcal{L} + \lambda\mathcal{R}$, where $\mathcal{L}$ coincides with the LSE functional. Because one wants the consistency constraint to be strictly enforced, a very high value for the regularization parameter (e.g., $\lambda = 1000$) has to be chosed. Actually, this means that the solution is found minimizing the regularity functional within the subspace of all consistent gradients fields. After the optimal gradient field is computed, the unwrapped phase is obtained by a path-independent integration.

Of particular interest is the "integer" version of this approach. The absolute phase gradients to estimate are written in terms of the observed phase gradients to which an integer number of $2\pi$ cycles is added. That is

$$\begin{aligned} \nabla_x f(i,j) &= \mathcal{W}[\nabla_x g(i,j)] + 2\pi k_x(i,j) f \\ \nabla_y f(i,j) &= \mathcal{W}[\nabla_y g(i,j)] + 2\pi k_y(i,j). \end{aligned} \tag{23}$$

The quantities $\nabla_x f(i,j)$ and $\nabla_y f(i,j)$ cannot assume any value except for only one of the values obtained through the previous relation. Now, the field to be estimated is $\{k_x, k_y\}$. From these values, the gradients of the absolute phase are calculated using (23). Two PU algorithms based on this scheme

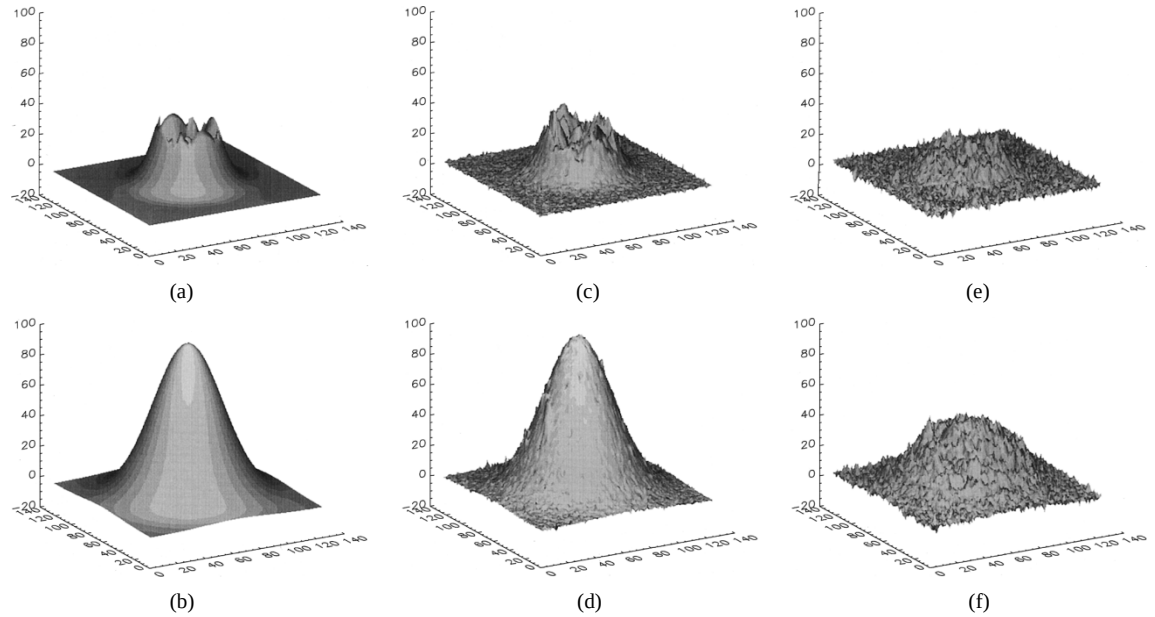


Fig. 5. First-order regularization solutions, through residue analysis, relative to the phase signals depicted in Fig. 1. Aliased phase signal: (a) noiseless, (c) noise with $\gamma = 0.9$, (e) noise with $\gamma = 0.6$. Phase signal without aliasing: (b) noiseless, (d) noise with $\gamma = 0.9$, (f) noise with $\gamma = 0.6$.

were independently proposed by Costantini [24] and Flynn [25]. The corresponding functional is

$$\mathcal{L} = \sum_{(i,j)\in I_x} |k_x(i,j)| + \sum_{(i,j)\in I_y} |k_y(i,j)| \tag{24}$$

subject to the constraint

$$\mathcal{R} = \sum_{i,j} [k_x(i,j) + k_y(i+1,j) - k_x(i,j+1) - k_y(i,j)]^2. \tag{25}$$

It is obtained by (18) and (23) and must be zero. In Bayesian formalism, it is modeled by two level pdf. The *a priori* assumption enters through the residue analysis, which states that the *a priori* probability distribution is zero everywhere except for the configurations satisfying the consistency constraint (25). The functional (24) is similar to (21) minimized by the LSE algorithm. It must be noted that this integer algorithm states that the solution $\{k_x, k_y\}$ has to correct for all residues, thereby producing a consistent field. Residues are no longer only a drawback; they help the PU. These algorithms automatically locate the branch cuts, which are lines connecting opposite sign residues. Across the branch cuts, the absolute phase gradients exceed $\pi$. As a consequence, the branch cuts must be avoided when integrating the estimated absolute phase gradients in order to have a consistent phase field. Many *ad hoc* attempts were made to identify these branch cuts [22], [51], [52]. Costantini's and Flynn's algorithm solutions give the branch cuts network across which the total absolute phase discontinuity is minimum. The results obtained using Costantini's algorithm are depicted in Fig. 5. It must be noted that Costantini's and Flynn's algorithms do not filter out the noise unless it produces residues. As can be seen, the results obtained with the nonaliased noisy phase signals are slightly better than that produced by the LSE approach. However, the same figure makes it evident how these algorithms share the same trouble characterizing the LSE solution when they are applied to aliased phase signals. This happens because the adopted *a priori* assumption is always that the adjacent absolute phase samples differ less than $\pi$.

### *D. Bayesian Estimation through Residue Analysis: Second-Order Regularization*

Recently, a new algorithm has been developed for solving PU in the framework of Bayesian inference [26]. The functional which is minimized comes from the prior model (16) with phase derivatives written using (23)

$$\begin{aligned}\mathcal{P}(f\,|\,g) = &\sum_{(i,j)\in I_{xx}} \{\mathcal{W}[\nabla_x g(i+1,j)] + 2\pi k_x(i+1,j) \\ &- \mathcal{W}[\nabla_x g(i,j)] - 2\pi k_x(i,j)\}^2 \\ &+ \sum_{(i,j)\in I_{xy}} \{\mathcal{W}[\nabla_y g(i+1,j)] + 2\pi k_y(i+1,j) \\ &- \mathcal{W}[\nabla_y g(i,j)] - 2\pi k_y(i,j)\}^2 \\ &+ \sum_{(i,j)\in I_{yx}} \{\mathcal{W}[\nabla_x g(i,j+1)] + 2\pi k_x(i,j+1) \\ &- \mathcal{W}[\nabla_y g(i,j)] - 2\pi k_y(i,j)\}^2 \\ &+ \sum_{(i,j)\in I_{yy}} \{\mathcal{W}[\nabla_y g(i,j+1)] + 2\pi k_y(i,j+1) \\ &- \mathcal{W}[\nabla_y g(i,j)] - 2\pi k_y(i,j)\}^2 \end{aligned} \tag{26}$$

subject to (25). The sets used are subsets of $I$ defined in such a way that all second derivatives can be defined. The constraint is the same as that used in [24] and expresses the condition that all residues are "absorbed" by the solution.

The functional (26) takes into account both the data and the estimated integer $2\pi$ cycles to adjoin to the data (in order to calculate the first derivatives to be summed along the smallest closed path). It should be noted the $k$-field estimate enables the reconstruction of absolute phase gradients and, by integration, of the phase field $f$. As a consequence, the notation $\mathcal{P}(f\,|\,g)$ is justified even if what we actually estimate is $\{k_x, k_y\}$. The peculiarity of this functional is the missing of the LSE term.

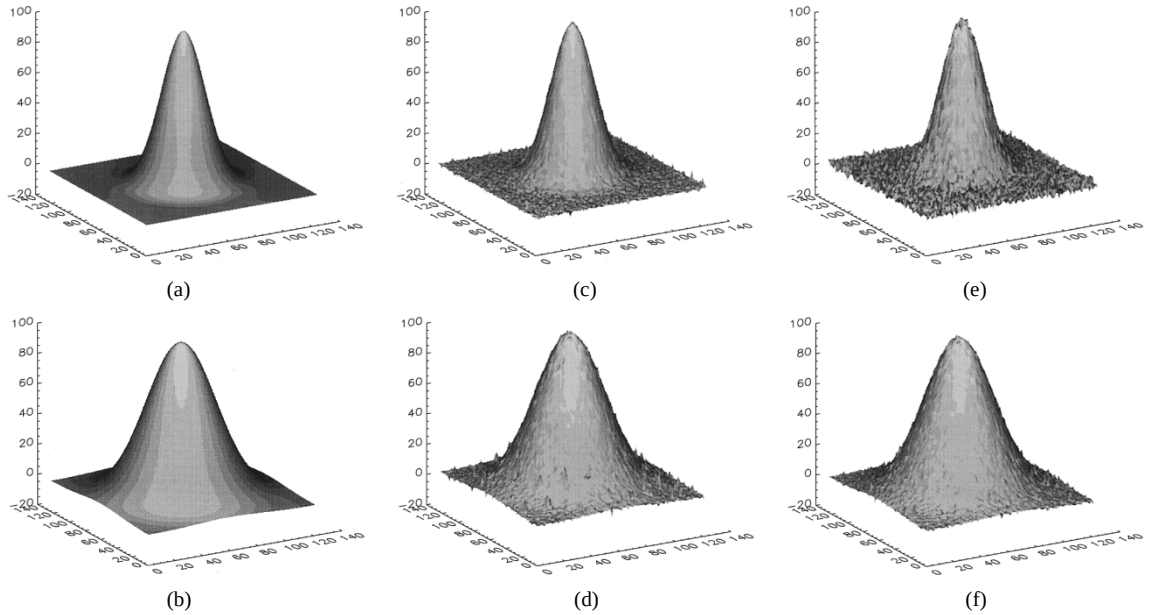


Fig. 6. Second-order regularization solutions, through residue analysis, relative to the phase signals depicted in Fig. 1. Aliased phase signal: (a) Noiseless, (c) noise with $\gamma = 0.9$, (e) noise with $\gamma = 0.6$. Phase signal without aliasing: (b) Noiseless, (d) noise with $\gamma = 0.9$, (f) noise with $\gamma = 0.6$.

This means that we avoid fitting $\{\nabla f(i,j)\}$ to $\{\mathcal{W}[\nabla g(i,j)]\}$, thereby releasing (4). In doing this, we take more *a priori* information on the nature of the sampling process. This information comes from the residue map, which highlights the fact that the sampled phase signal from which the interferogram is extracted is aliased. In these conditions, the use of (5) is not justified. All distortions in solutions obtained from the previous algorithms come from the abuse of (5). Condition (4) is substituted by the simple assumption on the regularity of the first derivatives $\{\nabla f(i,j)\}$. It is expressed through the minimization of the phase Laplacian. This way of writing the regularization functional avoids the problem of biased solutions encountered in all previous algorithms. This is due to the fact that the *a priori* assumption (4) is released. It is substituted with the less restrictive one that the absolute phase gradients in adjacent samples tend to be as likely as possible. In Fig. 6, the results obtained applying this algorithm on the testing interferograms are reported. In principle, this algorithm allows the reconstruction of phase surfaces having an arbitrary slope. Moreover, this algorithm does not filter out noise, except that which creates residues. We must note that the results of unwrapping are much better than those of LSE.

### *E. Maximum Entropy Regularization*

In this section, we describe an algorithm that gives an estimate of the unwrapped phase using the *maximum entropy principle* (19) to regularize the PU [34][1]

$$\mathcal{P} = \mathcal{L}_{\mathrm{LSE}} + \lambda \sum_{(i,j)\in I} f(i,j) \ln f(i,j). \tag{27}$$

The likelihood is still expressed through the LSE functional. Instead, the regularity is posed in a statistical manner [see (19)]. It is stated that in absence of the constraint due to observations, the absolute phase would assume the same value everywhere because this is the configuration having the minimum information and, therefore, the maximum entropy. The knowledge we have from data reduces the entropy of the absolute phase surface and gives it the shape we are looking for. The results of this algorithm are very similar (in tested conditions) to those obtained using a first-order regularization; see Section IV-B and Fig. 4(c) and (d). For example, the original aliased phase signal profile and the corresponding solution obtained by means of this algorithm are depicted in Fig. 8. The data refer to a noisy phase signal with $\gamma = 0.9$. The solution has been obtained using the parameter $\lambda = 0.1$.

### *F. Bayesian Model-Based Approach to PU: Fractal Model*

In any of the previous algorithms, the attention was on the real absolute phase field $\{f(i,j)\}$ to be reconstructed from the estimated absolute phase derivatives $\nabla f(i,j)$. In this section, an approach that addresses the solution of PU by using directly the complex phase signal $\exp(\imath f)$ is described. An *a priori* assumption is made on the physical quantity (e.g., height) to be reconstructed instead of the sampling process. These algorithms assume the absolute phase signal $f$ to be described by a model which can be a 2-D polynomial model [53], [54], Gaussian Markov random field model [35], or a fractal model [36]. Here, we present the latter method.

According to Bayesian modeling, we must to define the likelihood for the signal and the prior for the solution (see Section III). The likelihood term characterizes the loss of information in the observations due to the measurement noise and explicitly contains the prior knowledge on the data acquisition system

$$p(g \mid f) = p_n(\mathcal{D}(\mathcal{I}(g), \mathcal{J}(f))) \tag{28}$$

where $p_n$ is the probability density function (p.d.f.) of the noise, $\mathcal{J}(f) = \exp(\imath(f_1 - f_2)) = \exp(\imath f)$ characterizes the forward

[1]Both $f$ and $g$ have to be properly normalized in order to fulfill the condition $\sum f = 1$.

model or interferogram generation process, and $\mathcal{I}(g) = \exp(\imath g)$ is the observed complex interferogram (see Section II). $\mathcal{D}$ represents a distance measure defined as

$$\begin{aligned}\mathcal{D}(\omega) &= \mathcal{D}(\mathcal{I}(g), \mathcal{J}(f)) \\ &= \sum_{(\alpha,\beta)\in\Omega} \eta(\omega) \cdot \|\mathcal{F}\{\mathcal{I}(g) \cdot \mathcal{J}^*(f)\}(\omega)\|^2 \end{aligned} \quad (29)$$

where

- $\mathcal{F}$ Fourier transform operator;
- $\omega$ frequency;
- $\eta(\omega)$ quadratic function of frequency;

and the summation is done for all samples in the domain $\Omega$. The distance $\mathcal{D}(\omega)$ is a comparison of the observed complex valued interferogram $\mathcal{I}(g)$ with a *simulated* interferogram $\mathcal{J}(f)$, assuming a certain solution $f$ for the absolute phase. This distance corresponds to likelihood term $\mathcal{L}$ in (10). The role of the function $\eta(\omega)$ is to filter the low-frequency component emphasizing the comparison at high frequencies.

Relying on the observation that relief conserves a certain statistical characteristics over a wide range of scales, several techniques have been developed to perform a multiscale characterization of the elevation data. The statistics of terrain surface has been shown to agree, at least in a limited range of scales, with that of specific fractal models. Several attempts have been made to characterize digital elevation models (DEMs) by means of the fractal theory [55], [56]. The most popular model used in this respect is the fractional Brownian motion (fBm) model [57]. As we have seen already in the last part of the Section III-C, the increments of fBm process have interesting properties. They are strict sense stationary and additionally self similar, i.e., its increments verify $B_H(t+s) - B_H(t) = \alpha^{-H}(B_H(t+\alpha s) - B_H(t))$, where "$=$" means identical distributions, and $\alpha$ is a scale factor. Based on these properties, the direct estimation of the fractal process is substituted with the estimation of a stationary process: the fBm increments. The self similarity property is used to build a dyadic representation of the signal. The reconstruction of a fBm process is derived as midpoint deflections that generate a sequence of approximations at different scales. The dyadic partition is related to a tree representation of the recursive steps of our construction; thus, a multiscale stochastic process is defined. A random field $F$ is defined as having, as a support, the tree structure. The interscale causalities are described by the Markovian transition probability density functions [55]: $p(F^s \mid F^\sigma, \sigma < s) = p(F^s \mid F^{\gamma s}), F^s = (f_0^s, f_1^s, \ldots, f_p^s)$, where $s, \sigma$ are the scale indices, $\gamma$ is scale transition operator, and $p$ is the number of samples at a given scale. These models allow synthesis of finer scale $F^s$ of a random field beginning with its coarser scale $F^{s-1}$. The stochastic process represents the new information in a way similar to the detail signal in the wavelet transform domain. The models are inspired from the multigrid techniques in numerical analysis and define consistent multiscale Markov random fields starting from a single resolution. The estimation principle consists in solving a sequence of global optimization problems defined on a sequence of embedded configuration subspaces accepting constraints in form of prior distributions.

We use a *scene understanding technique*, that is, we try to find the scene that best explains the observed data. The problem statement is as follows: Minimize an error measure between the observed and the simulated data. The simulated data are iteratively obtained using the forward model as a function of the current height estimate. The terrain height $\hat{h} = (constant \cdot f)$ is obtained as a MAP estimator, having, as prior information, the fBm model. Although the fractal process is nonstationary, but knowing that its differential process is stationary and Gaussian, we generate the prior model using a midpoint displacement method, i.e., we sample from a stationary Gaussian process to generate the guess of the local terrain height increments at each scale. A modified likelihood is used in order to avoid the inversion of the wrapping operator: The actual complex interferogram $\mathcal{I}$ is compared with the simulated complex interferogram $\mathcal{J}(f) = \exp\{\imath f(\hat{h})\}$ in the Fourier transformed domain. The absolute phase $f(h)$ is proportional to $h$, which is the terrain height. The MAP estimate is obtained recursively in scale as

$$\begin{aligned}&\hat{h} = \arg\max_h \left\{ \sum_{(\alpha,\beta)\in\Omega} \eta(\omega) \cdot \|\mathcal{F}\{\mathcal{I}(g) \cdot \mathcal{J}^*(f)\}(\omega)\|^2 \right\} \\ &h \in \{\mathrm{fBm} \,|\, D = D_0\}. \end{aligned} \quad (30)$$

The MAP estimate is obtained using a simulated annealing algorithm. Sampling new values for $h$ from the posterior distribution is not possible in our case, and the prior has no closed form. Instead, we use a modified technique for optimization by separating the model and the likelihood terms.

1) Sample new values for $h$ from the prior distribution, i.e., make minor changes to the estimate determined by the model characteristics limiting the optimization space [this constraint [second line in (30)] corresponds to prior term $\mathcal{R}$ in (10)].
2) Evaluate the resulting change in the likelihood term using a Metropolis-like algorithm, thus minimizing the noise-dependent norm.

The process is iterative at one scale and recursive in scale. The used multiscale revisiting scheme has the role of updating the estimated values obtained at coarser scales as function of the ones obtained at finer scales.

In the algorithm, we used the likelihood p.d.f. approximated around its maximum with a Gaussian. This is the approximation that, in the case of a symmetric prior distribution, makes the MAP estimator equivalent to a least square estimator. The prior we used is it not symmetric. However, we observe that the quadratic weight function used in the definition of the likelihood is nothing but a differentiation in the time (space) domain:

$$\begin{aligned}\mathcal{D}(\omega) &= \sum a\omega^2 \cdot |\mathcal{F}\{\exp\{\imath(g-f)\}\}(\omega)|^2 \\ &= \sum \frac{a}{4\pi^2} \cdot |\imath 2\pi\omega \cdot \mathcal{F}\{\exp\{\imath(g-f)\}\}(\omega)|^2 \\ d(t) &= \sum \frac{a}{4\pi^2} \cdot |(\exp\{\imath(g-f)\})'(t)|^2 \\ &= \sum \frac{a}{4\pi^2} \cdot |(g'-f')(t)|^2 \end{aligned} \quad (31)$$

where $()'$ denotes the derivatives, and $a$ denotes a constant defining a shape of parabolic function $\eta$. The comparison of the observed interferogram with the simulated interferogram is equivalent to the comparison of the gradients of the observed and absolute phases.

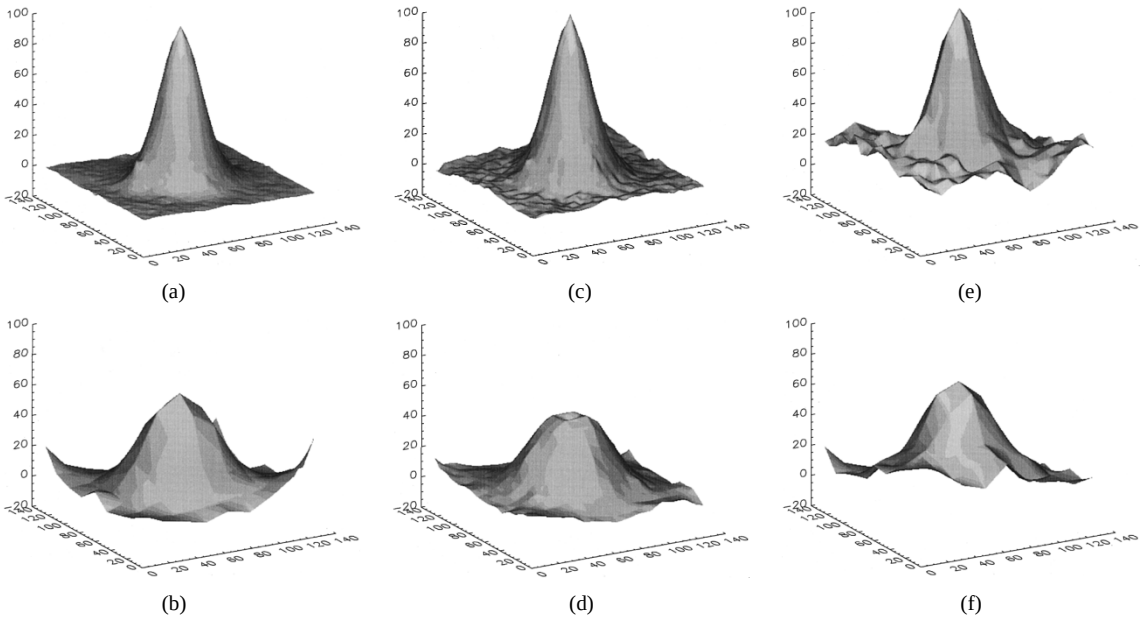



Fig. 7. Solutions obtained applying a model-based PU algorithm to the phase signals depicted in Fig. 1. The absolute phase signal is modeled by a fractal. Aliased phase signal: (a) Noiseless, (c) noise with $\gamma = 0.9$, (e) noise with $\gamma = 0.6$. Phase signal without aliasing: (b) Noiseless, (d) noise with $\gamma = 0.9$, (f) noise with $\gamma = 0.6$.

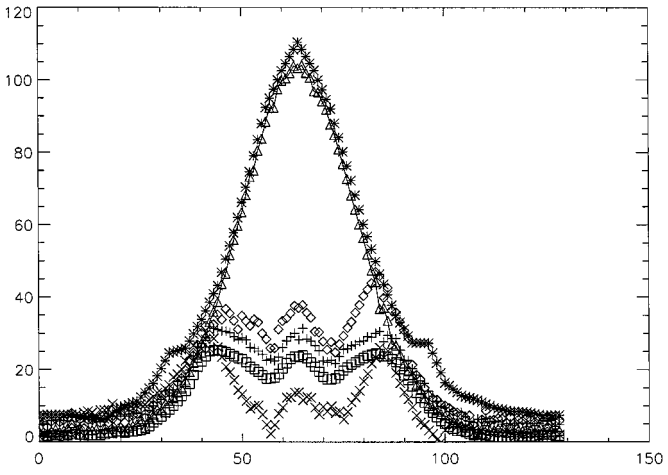


Fig. 8. Summary of the different PU schemes applied to the noisy aliased interferogram with $\gamma = 0.9$ shown in Fig. 1(e). The profiles refer to $(+)$ LSE, $(\square)$ first-order regularization, $(\diamond)$ first-order regularization with residue analysis and integer variables, $(\triangle)$ second-order regularization with residue analysis and integer variables, $(\times)$ maximum entropy regularization, $(*)$ fractal-based solution.

The infinite correlation of the assumed prior (the fBm) is a strong condition of surface regularity without restricting the solution to be smooth. Therefore, this PU scheme gives another way to release (4). The likelihood term is not measured using the reconstructed absolute phase slopes and the corresponding quantities observed in the interferogram. The results are depicted in Fig. 7 and are as promising as those of one of the previous method (see Section IV-D).

## V. Conclusion

In the preceding sections, we analyzed some of the PU schemes, emphasizing their *a priori* assumptions. We evaluated these schemes on synthetic interferograms with and without aliasing corrupted by different amounts of noise. The results relative to the noisy aliased phase signal with coherence $\gamma = 0.9$ are summarized in Fig. 8. It has been observed that all PU algorithms relying on the assumption that adjacent phase samples differ less than $\pi$ (as described in Sections IV-A, B, and C) can correctly solve the PU in the case of nonaliased phase signals with quite low noise. The solution efficiency and noise robustness of these algorithms can vary, depending on the adopted implementation. However, these algorithms cannot solve the severe problem of aliased phase signals, even in the absence of noise. They do not correctly reconstruct the trend of the absolute phase $f$. This behavior does not depend on the particular implementation. Indeed, this is a theoretical limit coming from (4) on the phase signal sampling these algorithms use.

The experimental results have also shown that the schemes described in Sections IV-D and F can solve PU both for nonaliased and aliased phase signals. In particular, two algorithms were analyzed: the first based on a second-order regularization [26] and the second on a fractal phase model [36]. These schemes seem to be more powerful because the *a priori* assumptions they use are less restrictive. Adjacent phase samples do not necessarily have to differ less than $\pi$. It suffices that their phase derivatives tend to be equal or their values fit a phase model. Up to now, many efforts were devoted to an efficient implementation of PU algorithms that can be cast into the LSE scheme. The considerations of this paper seem to suggest to concentrate these efforts on the potentially more powerful PU schemes described in Sections IV-D and F. This will give an efficient tool for solving PU in noisy interferograms where severe aliasing conditions occur.

## References

[1] *Special Issue J. Opt. Soc. Amer.*, vol. 67, no. 3, 1977.

[2] S. Nakadate and H. Saito, "Fringe scanning speckle-pattern interferometry," *Appl. Opt.*, vol. 24, pp. 2172–2180, 1985.

[3] D. L. Fried, "Least-squares fitting of a wave-front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Amer.*, vol. 67, pp. 370–374, 1977.

[4] R. H. Hudgin, "Wave-front reconstruction for compensated imaging," *J. Opt. Soc. Amer.*, vol. 67, pp. 375–378, 1977.

[5] R. J. Noll, "Phase estimates from slope-type wave-front sensors," *J. Opt. Soc. Amer.*, vol. 68, pp. 139–140, 1978.

[6] J. Hermann, "Least-squares wave front errors of minimum norm," *J. Opt. Soc. Amer. A*, vol. 70, no. 4, pp. 28–35, 1980.

[7] S. M. H. Song, S. Napel, N. J. Pelc, and G. H. Glover, "Phase unwrapping of MR phase images using poisson equation," *IEEE Trans. Image Processing*, vol. 4, pp. 667–676, Apr. 1995.

[8] A. Devaney, "Geophysical diffraction tomography," *IEEE Trans. Geosci. Remote Sensing*, vol. GE–22, pp. 3–13, 1984.

[9] ——, "Diffraction tomographic reconstruction from intensity data," *IEEE Trans. Signal Processing*, pp. 221–228, Jan. 1992.

[10] R. M. Goldstein and H. A. Zebker, "Interferometric radar mapping of ocean currents," *Nature*, vol. 328, pp. 707–709, 1987.

[11] D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, and T. Rabaute, "The displacement field of the Landers earthquake mapped by radar interferometry," *Nature*, vol. 364, pp. 138–142, 1993.

[12] S. R. Cloude and K. P. Papathanassiou, "Polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1551–1565, Oct. 1998.

[13] L. M. H. Ulander and P. O. Frölind, "Ultra-wideband SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1540–1550, Oct. 1998.

[14] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1975.

[15] H. P. Hjalmarson, L. A. Romero, D. C. Ghiglia, E. D. Jones, and C. B. Norris, "Extraction of phonon density-of-states from optical spectra," *Phys. Rev. B*, vol. B, no. 32, pp. 4300–4303, 1985.

[16] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping. Theory Algorithms and Software*. New York: Wiley, 1998.

[17] E. Rodriguez and J. M. Martin, "Theory and design of interferometric synthetic aperture radars," *Proc. Inst. Elect. Eng. F*, no. 139, pp. 147–159, 1992.

[18] D. Just and R. Bamler, "Phase statistics of interferograms with applications to synthetic aperture radar," *Appl. Opt.*, vol. 33, pp. 4361–4368, 1994.

[19] H. A. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 950–959, 1992.

[20] H. A. Zebker and R. M. Goldstein, "Topographic mapping from interferometric aperture radar observations," *J. Geophys. Res.*, vol. 91, no. B5, pp. 4993–4999, 1986.

[21] L. C. Graham, "Synthetic interferometric radar for topographic mapping," *Proc. IEEE*, vol. 62, pp. 763–768, 1974.

[22] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, no. 4, pp. 713–720, 1988.

[23] M. Costantini, A. Farina, and F. Zirilli, "A fast phase unwrapping algorithm for SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 452–460, Jan. 1999.

[24] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 813–821, May 1998.

[25] T. J. Flynn, "Two dimensional phase unwrapping with minimum weighted discontinuity," *J. Opt. Soc. Amer. A*, vol. 14, no. 10, pp. 2692–2701, 1997.

[26] L. Guerriero, G. Nico, G. Pasquariello, and S. Stramaglia, "A new regularization scheme for phase unwrapping," *Appl. Opt.*, vol. 37, no. 14, pp. 3053–3058, 1998.

[27] D. C. Ghiglia and L. A. Romero, "Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods," *J. Opt. Soc. Amer. A*, vol. 11, pp. 107–117, 1994.

[28] M. D. Pritt and J. S. Shipman, "Least-squares two-dimensional phase unwrapping using FFT's," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 706–708, 1994.

[29] M. D. Pritt, "Phase unwrapping by means of multigrid techniques for interferometric SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 728–738, 1996.

[30] G. Fornaro, G. Franceschetti, and R. Lanari, "Interferometric phase unwrapping using Green's formulation," *IEEE Trans. Geosci. Remote Sensing*, vol. 34, pp. 720–727, 1996.

[31] G. Fornaro, G. Franceschetti, R. Lanari, D. Rossi, and M. Tesauro, "Interferometric SAR phase unwrapping using the finite element method," *Proc. Inst. Elect. Eng., Radar Sonar Navig.*, vol. 144, pp. 266–274, 1997.

[32] I. Lyuboshenko and H. Maître, "Phase unwrapping for interferometric synthetic aperture radar by use of helmholtz equation eigenfunctions and first Green's identity," *J. Opt. Soc. Amer. A*, vol. 16, no. 2, pp. 378–395, 1999.

[33] J. Marroquin and M. Rivera, "Quadratic regularization functionals for phase unwrapping," *J. Opt. Soc. Amer. A*, vol. 12, no. 11, pp. 2393–2400, 1995.

[34] M. Datcu, "Maximum entropy solution for InSAR phase unwrapping," in *Proc. IGARSS*, 1996, pp. 310–314.

[35] J. M. N. Leitão and M. A. T. Figueiredo, "Absolute phase image reconstruction: A stochastic nonlinear filtering approach," *IEEE Trans. Image Processing*, vol. 7, pp. 868–882, June 1998.

[36] M. Datcu and G. Palubinskas, "Multiscale bayesian terrain height estimation from SAR interferometry," in *Proc. IGARSS*, 1998, pp. 88–90.

[37] H. D. Griffiths, "Interferometric synthetic aperture radar," *Electron. Commun. Eng. J.*, vol. 7, no. 6, pp. 247–256, 1995.

[38] K. Itoh, "Analysis of the phase unwrapping algorithm," *Appl. Opt.*, vol. 21, p. 2470, 1982.

[39] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inv. Prob.*, vol. 14, pp. R1–R54, 1998.

[40] W. Xu and I. Cumming, "A region-growing algorithm for InSAR phase unwrapping," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 124–134, Jan. 1999.

[41] R. T. Cox, "Probability, frequency and reasonable expectation," *Amer. J. Phys.*, vol. 17, pp. 1–13, 1946.

[42] T. Poggio, V. Torre, and C. Kock, "Computational vision and regularization theory," *Nature*, vol. 317, pp. 314–319, 1985.

[43] R. Bamler, N. Adam, G. W. Davidson, and D. Just, "Noise-induced slope distortion in 2-D phase unwrapping by linear estimators with application to SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 913–921, May 1998.

[44] B. R. Hunt, "Matrix formulation of the reconstruction of phase values from phase differences," *J. Opt. Soc. Amer.*, vol. 69, no. 3, pp. 393–399, 1979.

[45] H. Takajo and T. Takahashi, "Noniterative methods for obtaining the exact solution for the normal equation in least-squares phase estimation from phase difference," *J. Opt. Soc. Amer. A*, vol. 5, pp. 1818–1827, 1988.

[46] M. D. Pritt, "Congruence in least-squares phase unwrapping," in *Proc. IGARSS*, 1997, pp. 875–877.

[47] U. Spagnolini, "2-D phase unwrapping and istantaneous frequency estimation," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 579–589, 1995.

[48] E. Trouvé, J. M. Nicolas, and H. Maître, "Improving phase uwrapping by the use of local frequency estimates," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 1963–1972, May 1998.

[49] G. W. Davidson and R. Bamler, "Multiresolution phase unwrapping for SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 163–174, Jan. 1999.

[50] M. Rivera, J. Marroquin, M. Servin, and R. Rodriguez-Vera, "Fast algorithms for integrating inconsistent gradient fields," *Appl. Opt.*, vol. 36, pp. 8381–8390, 1997.

[51] J. R. Buckland, J. M. Huntley, and S. R. E. Turner, "Unwrapping noisy phase maps by use of a minimum-cost-matching algorithm," *Appl. Opt.*, vol. 34, pp. 5100–5108, 1995.

[52] J. A. Quiroga, A. Gonzales-Cano, and E. Bernabeu, "Stable-marriages algorithm for preprocessing phase maps with discontinuity sources," *Appl. Opt.*, vol. 34, pp. 5029–5038, 1995.

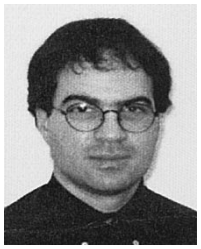
[53] Z.-P. Liang, "A model-based method for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 15, pp. 893–897, 1996.

[54] B. Friedlander and J. M. Francos, "Model based phase unwrapping for 2-D signals," *IEEE Trans. Signal Processing*, vol. 44, pp. 2999–3007, Dec. 1996.

[55] K. C. Clarke and D. M. Schweizer, "Measuring the fractal dimension of natural surfaces using a robust fractal estimator," *Cartogr. Geograph. Inform. Syst.*, vol. 18, no. 1, pp. 37–47, 1991.

[56] J. Huang and D. L. Turcotte, "Fractal image analysis: Application to the topography of Oregon and synthetic images," *J. Opt. Soc. Amer. A*, vol. 7, no. 6, pp. 1124–1130, 1990.

[57] G. W. Wornell, "Wavelet-based representations for the $1/f$ family of fractal processes," *Proc. IEEE*, vol. 81, pp. 1428–1450, 1993.

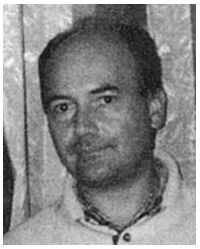


**Giovanni Nico** (M'99) received the Laurea and Ph.D. degrees in physics from the University of Bari, Bari, Italy, in 1993 and 1999, respectively. His doctoral dissertation was on SAR interferometry.

From November 1997 to June 1998, he visited the German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Oberpfaffenhofen, Germany. Currently, he is with the Space Applications Institute (SAI), Joint Research Centre of the European Commission, Ispra, Italy. His main research interests are in SAR interferometry and phase unwrapping.

**Gintautas Palubinskas** received the M.S. and Ph.D. degrees in mathematics from Vilnius University, Vilnius, Lithuania, in 1981 and Institute of Mathematics and Informatics (IMI), Vilnius, in 1991. His doctoral dissertation was on spatial image recognition.

He was a Research Scientist at the IMI from 1981 to 1997. Since 1997, he has been a Research Scientist at German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), Obepfaffenhofen, Germany. From 1993 to 1997, he was a Visiting Research Scientist at DFD/DLR; the Department of Geography, Swansea University, Swansea, U.K.; Institute of Navigation, Stuttgart University, Stuttgart Germany; and the Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany. His current interests are in image classification and fusion, Bayesian inference, SAR interferometry, and remote sensing.



**Mihai Datcu** received the Ph.D. degrees in electronics and telecommunications from the University "Politehnica," Bucharest, Romania (UPB), in 1986, and the title "Habilitation à diriger des recherches" from Université Louis Pasteur, Strasbourg, France, in 1999.

He has held a professorship in electronics and telecommunications at UPB since 1981. Currently, he is a Group Leader with the Remote Sensing Technology Institute (IMF) of the German Aerospace Center (DLR), Oberpfaffenhofen, Germany. He is developing algorithms for scene understanding from synthetic aperture radar (SAR) and interferometric SAR data and model-based methods for information retrieval, and he conducts research in information theoretical aspects and semantic representations in advanced communication systems. He held Visiting Professor appointments from 1991 to 1992 at the Department of Mathematics, University of Oviedo, Oviedo, Spain, and from 1992 to 1993, 1996 to 1998, and 2000, he was with the Swiss Federal Institute of Technology (ETH), Žurich, Switzerland. He was a Guest Scientist with the Swiss Center for Scientific Computing (CSCS), Manno, in 1994. He was teaching stochastic image analysis, fractal analysis, image processing in medical sciences, and designing and developing new concepts and systems for image information mining, realistic visualization, query by image content from very large image archives, and new algorithms for parameter estimation. His current interests are in Bayesian inference, information theory, stochastic processes, model-based scene understanding, and image information mining, with applications in SAR and optical imaging.