

# Absolute Phase Image Reconstruction: A Stochastic Nonlinear Filtering Approach

José M. N. Leitão, *Member, IEEE*, and Mário A. T. Figueiredo, *Member, IEEE*

**Abstract**— This paper formulates and proposes solutions to the problem of estimating/reconstructing the absolute (not simply modulo- $2\pi$ ) phase of a complex random field from noisy observations of its real and imaginary parts. This problem is representative of a class of important imaging techniques such as interferometric synthetic aperture radar, optical interferometry, magnetic resonance imaging, and diffraction tomography. We follow a Bayesian approach; then, not only a probabilistic model of the observation mechanism, but also prior knowledge concerning the (phase) image to be reconstructed, are needed. We take as prior a nonsymmetrical half plane autoregressive (NSHP AR) Gauss–Markov random field (GMRF). Based on a reduced order state-space formulation of the (linear) NSHP AR model and on the (nonlinear) observation mechanism, a recursive stochastic nonlinear filter is derived. The corresponding estimates are compared with those obtained by the extended Kalman–Bucy filter, a classical linearizing approach to the same problem. A set of examples illustrate the effectiveness of the proposed approach.

**Index Terms**— Absolute phase imaging, Bayesian estimation, image reconstruction, interferometric imaging, Kullback–Leibler divergence, nonlinear filtering, phase unwrapping, stochastic filtering, 2-D Kalman–Bucy filtering.

## I. INTRODUCTION

### A. Absolute Phase Imaging and Its Applications

CONSIDER the conceptual imaging model depicted in Fig. 1, in which a discrete image/surface,  $\mathbf{x} = \{x_{i,j} \in \mathbb{R}, i = 0, \dots, M-1, j = 0, \dots, N-1\}$ , is observed through a pair of nonlinear (sine and cosine) functions corrupted by additive white Gaussian noise. This paper addresses the nonlinear image reconstruction problem that consists in estimating  $\mathbf{x}$  from this pair of nonlinear noisy observations,  $\mathbf{y}^c$  and  $\mathbf{y}^s$ . Interpreting the observed images,  $\mathbf{y}^c$  and  $\mathbf{y}^s$ , as noisy versions of the real and imaginary parts of a complex image, the problem can also be stated as that of estimating the (absolute) phase of this complex image. This scenario captures the essential common features of a general class of imaging/sensing techniques where some physical quantity is inferred from absolute (i.e., not simply modulo- $2\pi$ ) phase measurements [39]. Following are examples of such techniques.

Manuscript received July 19, 1996; revised January 20, 1997. This work was supported in part by the PRAXIS XXI Program of the Portuguese Ministry of Science and Technology under Project 2/2.1/TIT/1580/95. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. W. Clem Karl.

The authors are with the Instituto de Telecomunicações and Departamento de Engenharia Electrotécnica e de Computadores, Instituto Superior Técnico, 1096 Lisboa Codex, Portugal (e-mail: mtf@lx.it.pt).

Publisher Item Identifier S 1057-7149(98)03993-1.

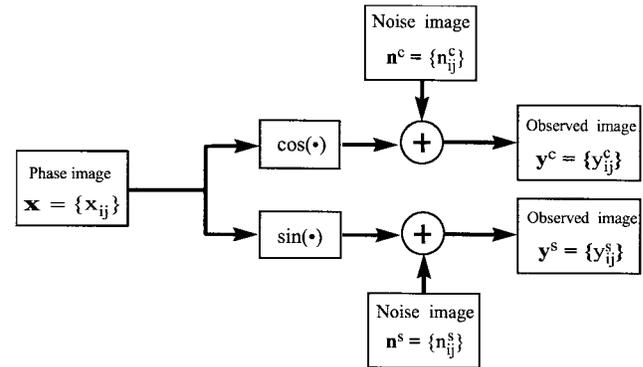


Fig. 1. Observation model.

- *Interferometric synthetic aperture radar* (InSAR), where phase measurements are used to produce topographic maps. More precisely, (absolute) phase differences between complex SAR images acquired by different antennas [24], [26], [60], or by the same antenna at separate passes [21], [25], [27], are a function of the elevation of the observed terrain; see reviews in [1] and [17], and references therein.
- *Optical interferometry*, where (absolute) phase differences are used to obtain information, such as shape, displacement, or vibration, of a surface under inspection; see, e.g., [28], [51] and [57].
- *Magnetic resonance imaging* (MRI), in which absolute phase measurements allow increasing the dynamic range of phase contrast velocity images; references [4], [12], [30], and [54] consider this problem.
- *Diffraction tomography* (e.g. geophysical tomography or ultrasound medical tomography), where the Rytov approximation yields a mapping between the observed object and the absolute phase of the measured field; for details, see [14], [15], and [34].

### B. The “Phase Unwrapping” Approach

Conventional approaches to the problem of obtaining absolute phase images follow a two-step procedure:

- Step 1)** determination of modulo  $2\pi$  phase values, the so-called *interferogram* or *wrapped phase* image;
- Step 2)** *phase unwrapping* (i.e., determination of absolute values from modulo  $2\pi$  ones) supported on some heuristic or *ad hoc* phase continuity criterion.

This *phase unwrapping* approach has been applied under several forms to all of the above mentioned imaging modalities.

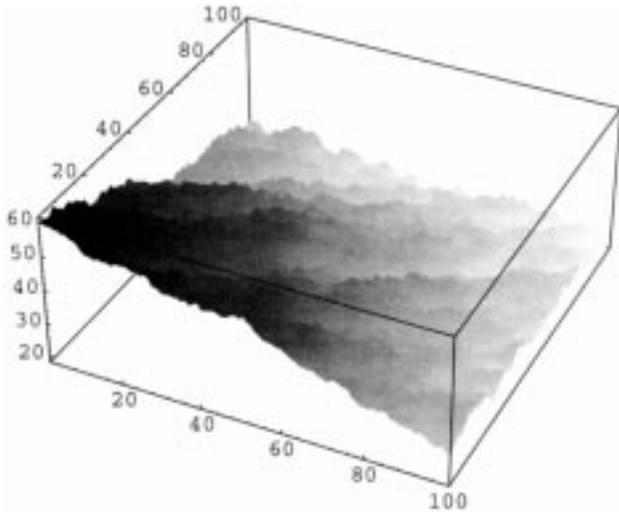


Fig. 2. Original phase surface/image (example). The surface height values are also gray-level coded.

References [20], [25], [29], [36], [42], [43], [45], [52], and [55] describe applications to InSAR. Phase unwrapping techniques for MRI have been presented in [4], [12], [30], and [54]. In the field of optical interferometry, several methods were published in [28], [46], [47], [51], and [57]. Reference [34] includes a phase unwrapping algorithm in the context of diffraction tomography.

### C. The Proposed Approach

Phase estimation has been previously studied within the stochastic nonlinear filtering (NLF) framework for (one-dimensional) time signals [8], [9], [48]. In particular, recursive absolute phase estimation algorithms have been studied in [40], [41], and [49]. In this paper, we propose a stochastic NLF approach to the problem of estimating/reconstructing absolute phase (two-dimensional) images; it stems from recognizing the fact that this is a two-dimensional (2-D) version of the problem considered in [40], [41], and [49]. Being essentially a Bayesian approach to the estimation of a random process/field from its noisy observations, stochastic NLF is supported, not only on probabilistic models of the observation mechanism, but also on prior knowledge concerning the (phase) image to be reconstructed.

Markov random fields (MRF) have been widely used as (Bayesian) priors in many image processing and computer vision problems; [6], [10], [13], [16], [23], [50], [58] are a few relevant references on this subject. In this paper, the original phase image is modeled as a sample of a Gauss–Markov random field (GMRF) [10], [11], [13], [16], [58]. More specifically, and having recursive filtering in mind, we consider causal nonsymmetrical half plane (NSHP) autoregressive (AR) GMRF’s; this class of models, formalizing prior knowledge in a probabilistic way, allows building recursive filtering algorithms.

A recursive stochastic filtering approach requires a state-space formulation of the NSHP-AR model. Here, we adopt a *reduced order model* (ROM) similar to the one proposed in

[2], [3], and [33], to build ROM Kalman filters (ROMKF) for problems with linear observation models.

Supported on the ROM, and adopting stochastic nonlinear filtering methodologies proposed in [40] and [41], we introduce a recursive scheme that estimates the absolute phase image directly from its noisy sine and cosine observations. Accordingly, our technique should not be classified as a conventional phase-unwrapping algorithm since it does not start from wrapped phase data.

The *extended Kalman–Bucy filter* (EKBF) [8], [22], [32], which is a classical solution to nonlinear stochastic filtering problems, is also considered and compared with the developed NLF algorithm.

Noncausal MRF priors could also be used, as in [5] and [50]; there, a recursive formulation of noncausal MRF’s is proposed and image restoration/reconstruction is interpreted as a fixed interval smoothing problem (with linear observations) solved by double-sweep linear recursive filtering techniques. A main issue with the nonlinear observation model herein considered would be how to solve fixed interval smoothing problems by applying the type of nonlinear algorithms presented in the sequel.

### D. Paper Overview

The next section formally describes the observation model considered. Section III introduces the adopted prior model. The proposed stochastic NLF algorithm and the EKBF are described in Section IV. Section V presents experimental results and Section VI ends the paper with some concluding remarks.

## II. OBSERVATION MODEL

The structure of the observation model is depicted in Fig. 1. Let

$$\mathbf{x} = \{x_{i,j} \in \mathbb{R}, i = 0, \dots, M-1, j = 0, \dots, N-1\} \quad (1)$$

be the  $\mathbb{R}^{MN}$ -valued,  $M \times N$  original phase image. Let the observations be denoted by

$$\mathbf{y}^c = \{y_{i,j}^c, i = 1, \dots, M, j = 1, \dots, N\} \quad (2)$$

and

$$\mathbf{y}^s = \{y_{i,j}^s, i = 1, \dots, M, j = 1, \dots, N\} \quad (3)$$

the so-called *in-phase* (cosine) and *quadrature* (sine) images (associated to the phase field  $\mathbf{x}$ ), which are additively corrupted by the independent white Gaussian noise fields

$$\mathbf{n}^c = \{n_{i,j}^c, i = 0, \dots, M-1, j = 0, \dots, N-1\} \quad (4)$$

and

$$\mathbf{n}^s = \{n_{i,j}^s, i = 0, \dots, M-1, j = 0, \dots, N-1\}. \quad (5)$$

Formally, the observation model is

$$\mathbf{y}_{i,j} = \begin{bmatrix} y_{i,j}^c \\ y_{i,j}^s \end{bmatrix} = \begin{bmatrix} \cos(x_{i,j}) \\ \sin(x_{i,j}) \end{bmatrix} + \begin{bmatrix} n_{i,j}^c \\ n_{i,j}^s \end{bmatrix},$$

for  $i = 0, \dots, M-1$ , and  $j = 0, \dots, N-1$ . (6)

We assume that the noise fields can be spatially variant; thus, we write  $\sigma_{ij}^2$  as the variance of the noise random

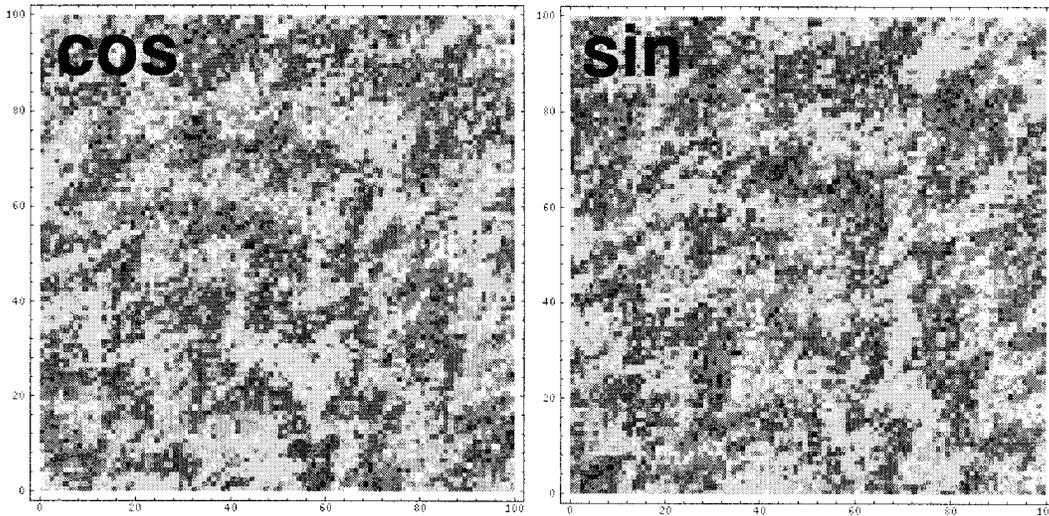


Fig. 3. Noisy ( $\sigma = 0.5$ ) cosine and sine observations, of the image/surface of Fig. 2.

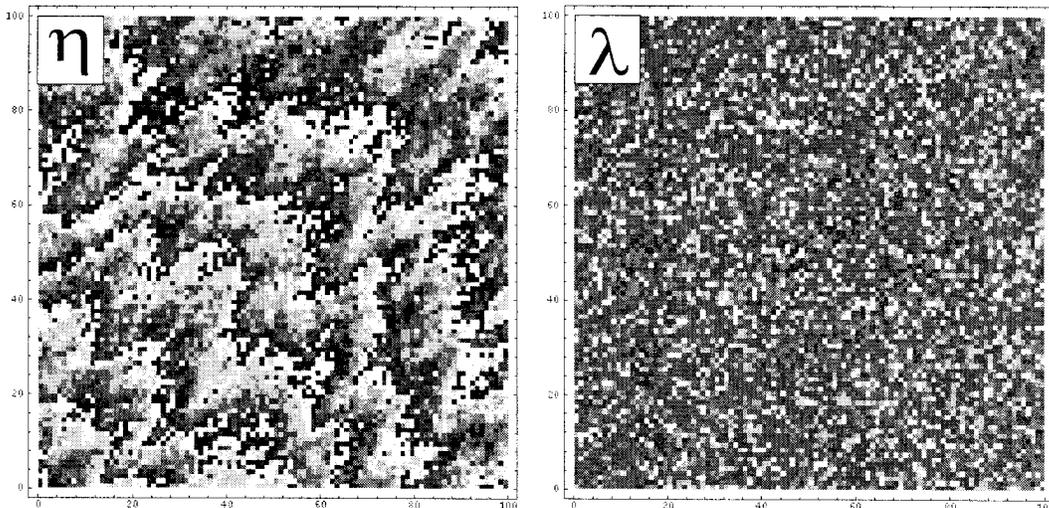


Fig. 4.  $\eta_{i,j}$ 's (interferogram) and  $\lambda_{i,j}$ 's [see (10) and (11)] obtained from the observations shown in Fig. 3.

variables associated to pixel  $(i, j)$ ,  $n_{i,j}^c$  and  $n_{i,j}^s$ , for  $i = 0, \dots, M-1$  and  $j = 0, \dots, N-1$ . Finally, we mention that, particularly in InSAR and optical interferometry, the presence of additive white Gaussian noise in the in-phase and quadrature components is in fact the commonly adopted model [31], [38], [44], [53]. As an example of data produced by this observation model, consider the phase surface/image of Fig. 2; the corresponding observations  $\mathbf{y}^c$  and  $\mathbf{y}^s$ , with homogeneous noise ( $\sigma_{i,j} = \sigma = 0.5$ ), are presented in Fig. 3.

From (6), and invoking the Gaussian nature of the noise, we obtain the pixelwise conditional probability density function of the pair of observations associated to  $x_{i,j}$ , as

$$p(y_{i,j}^c, y_{i,j}^s | x_{i,j}) = \frac{1}{2\pi\sigma_{i,j}^2} \exp \left\{ -\frac{1}{2\sigma_{i,j}^2} \left[ (y_{i,j}^c - \cos(x_{i,j}))^2 + (y_{i,j}^s - \sin(x_{i,j}))^2 \right] \right\} \quad (7)$$

which can be rewritten as

$$p(y_{i,j}^c, y_{i,j}^s | x_{i,j}) = \frac{\alpha_{i,j}}{2\pi\sigma_{i,j}^2} \exp\{\lambda_{i,j} \cos(x_{i,j} - \eta_{i,j})\} \quad (8)$$

where

$$\alpha_{i,j} = \exp \left\{ -\frac{1}{2\sigma_{i,j}^2} [1 + (y_{i,j}^c)^2 + (y_{i,j}^s)^2] \right\} \quad (9)$$

$$\lambda_{i,j} = \frac{1}{\sigma_{i,j}^2} \sqrt{(y_{i,j}^c)^2 + (y_{i,j}^s)^2} \quad (10)$$

$$\eta_{i,j} = \arctan \left( \frac{y_{i,j}^s}{y_{i,j}^c} \right) \in [0, 2\pi[. \quad (11)$$

Fig. 4 displays the  $\eta_{i,j}$ 's and  $\lambda_{i,j}$ 's obtained from the noisy data of Fig. 3.

Notice that  $p(y_{i,j}^c, y_{i,j}^s | x_{i,j})$  is the likelihood function of  $x_{i,j}$ , given the observations  $y_{i,j}^c$  and  $y_{i,j}^s$ . Then, it is clear from (8) that  $\{\eta_{i,j} + 2\pi l_{i,j}, l_{i,j} \in \mathcal{Z}\}$  is the (infinite) set of maximum likelihood (ML) estimates of  $x_{i,j}$ , all yielding the

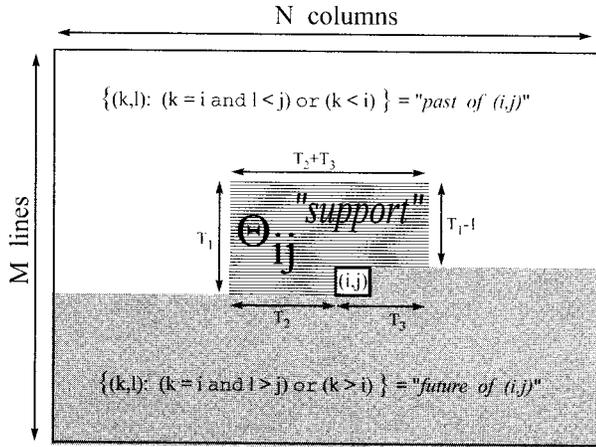


Fig. 5. NSHP Markov model.

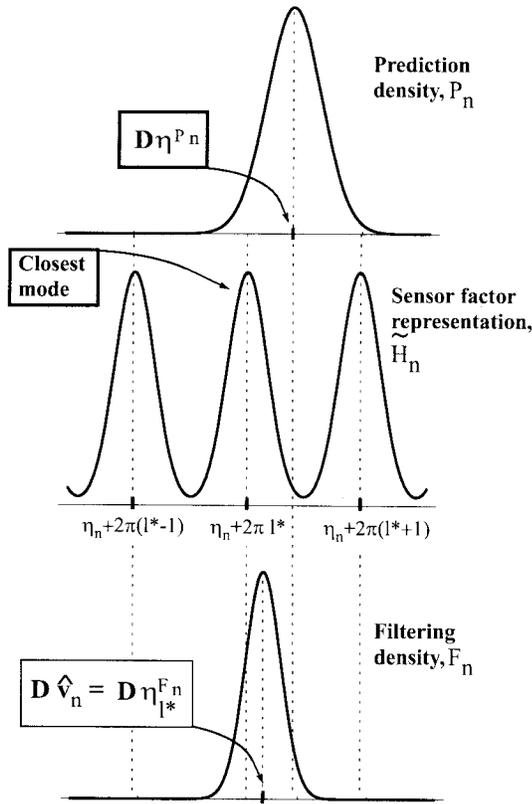


Fig. 6. Computation of the filtering density (see Section IV-F).

same maximum value of the likelihood function (observe that  $\sigma_{i,j}^2$ ,  $\alpha_{i,j}$ ,  $\lambda_{i,j}$ , and  $\eta_{i,j}$ , appearing in (8), are all independent of  $x_{i,j}$ ). Conventional phase unwrapping techniques, as the ones mentioned in Section I-B, usually take the  $\eta_{i,j}$ 's (i.e., the interferogram) as the observations and try to determine the  $l_{i,j}$ 's by imposing some continuity criterion.

Few approaches take into account, formally, the noisy nature of the observations. References [46] and [47] consider the observations as wrapped versions of the absolute phase, with additive white Gaussian noise; this is an invalid assumption [31], [38], [53]. The model based approach in [20] adopts an observation model similar to ours, i.e., real and imaginary

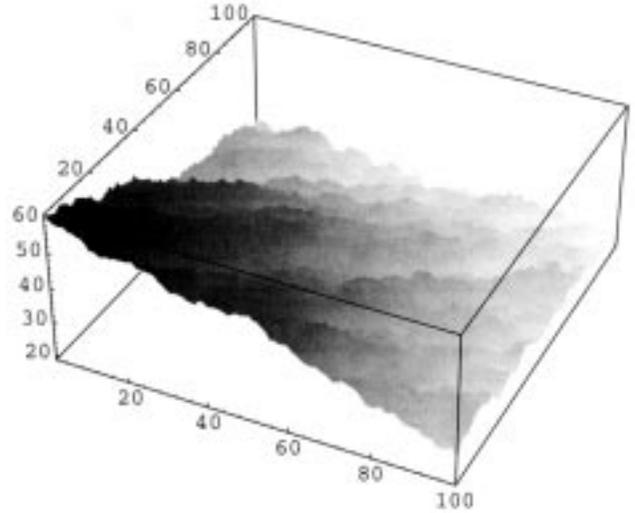


Fig. 7. Phase image/surface estimate obtained by the NLF from the observations shown in Fig. 3 (compare with Fig. 2).

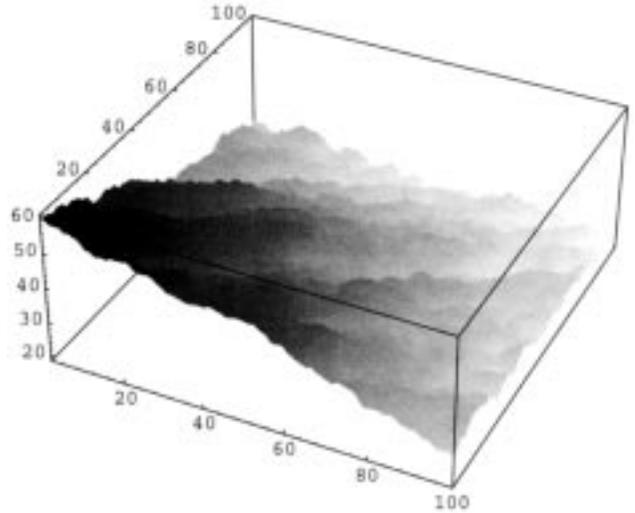


Fig. 8. Phase image/surface estimate obtained by the EKBF from the observations shown in Fig. 3 (compare with Fig. 2).

parts of a complex field with additive white Gaussian noise. However, that work differs considerably from ours: instead of a sample of a random field, it considers the original absolute phase image/surface as a 2-D polynomial whose parameters are to be estimated from the noisy data.

### III. ORIGINAL ABSOLUTE PHASE IMAGE MODEL

#### A. AR Gauss–Markov Model

We model the original phase image/surface as a sample of a causal, NSHP, AR GMRF [10], [13], [59]; specifically

$$x_{i,j} = \sum_{(k,l) \in \Theta_{i,j}} \theta_{i,j,k,l} x_{k,l} + u_{i,j} \quad (12)$$

where the  $u_{i,j}$ 's are independent and identically distributed (i.i.d.) zero-mean Gaussian variables of variance  $\mu^2$ , and  $\Theta_{i,j}$  (see Fig. 5) is the *support* of pixel  $(i,j)$  [13], [59]. Horizontal

(line-by-line), left to right, top to bottom recursion order (which coincides with the lexicographical order) is herein assumed.

When the position of the support is such that it requires pixels from outside the image, some assumption has to be made about those pixels. Here, we adopt *free boundary conditions* which, in the AR formulation, are equivalent to considering that the pixels near the boundaries simply have a smaller support. The model coefficients associated to *external* pixels are set to zero, so that no boundary values have to be defined, i.e.,

$$(k < 1 \text{ or } k > M \text{ or } l < 1 \text{ or } l > N) \Rightarrow \theta_{i,j,k,l} = 0. \quad (13)$$

For example, the first line of the image is taken as a one-dimensional (1-D) AR process.

The surface presented in Fig. 2 is a sample of a NSHP AR GMRF; its generating model (with the standard deviation of the driving noise set to  $\mu = 0.7$ ) is

$$x_{i,j} = 0.495x_{i,j-1} + 0.495x_{i-1,j} + 0.005x_{i-1,j-1} + u_{i,j}. \quad (14)$$

### B. State Space Formulation

For recursive stochastic filtering purposes, a state-space formulation of (12) is necessary. Adopting the procedure proposed in [59], the state vector would have to include several complete lines of pixels together with some boundary values; this approach would lead to a huge state vector of which only a few elements are really important. Alternatively, we adopt a *reduced order model* (ROM), as proposed in [2] and [3], where the state vector contains only the pixels in the support of the AR model.

Since we are assuming a lexicographical recursion order, the image pixels can be addressed by a single index  $n$ , which is related to the line and column coordinates by

$$x_n \equiv x_{l(n),c(n)} \quad (15)$$

where  $l(n)$  and  $c(n)$  are the *line* and *column* functions

$$l(n) = \left\lfloor \frac{n}{N} \right\rfloor \quad (16)$$

$$c(n) = n - \left\lfloor \frac{n}{N} \right\rfloor N \quad (17)$$

for  $n = 1, 2, \dots, MN$ . In (16) and (17),  $\lfloor a \rfloor$  stands for the *greatest integer smaller than a*. Notice that, if  $c(n) \neq N$ , then  $x_{n+1} = x_{l(n),c(n)+1}$ , and, if  $c(n) = N$ , then  $x_{n+1} \equiv x_{l(n)+1,1}$ .

The (reduced order) state vector  $\mathbf{v}_n$ , corresponding to pixel  $x_{l(n),c(n)}$ , has to contain all the “past” necessary to predict pixel  $x_{l(n),c(n)+1}$ , i.e., all the pixels in the support  $\Theta_{l(n),c(n)+1}$ ,

$$\mathbf{v}_n = \begin{bmatrix} x_{l(n),c(n)} & x_{l(n),c(n)-1} & \cdots & x_{l(n),c(n)-(T_2-1)} \\ x_{l(n)-1,c(n)-(T_2-1)} & \cdots & x_{l(n)-1,c(n)} \\ x_{l(n)-1,c(n)+1} & \cdots & x_{l(n)-1,c(n)+T_3} \\ \vdots & & & \\ x_{l(n)-(T_1-1),c(n)-(T_2-1)} & \cdots & x_{l(n)-(T_1-1),c(n)} \\ x_{l(n)-(T_1-1),c(n)+1} & \cdots & x_{l(n)-(T_1-1),c(n)+T_3} \end{bmatrix}^T$$

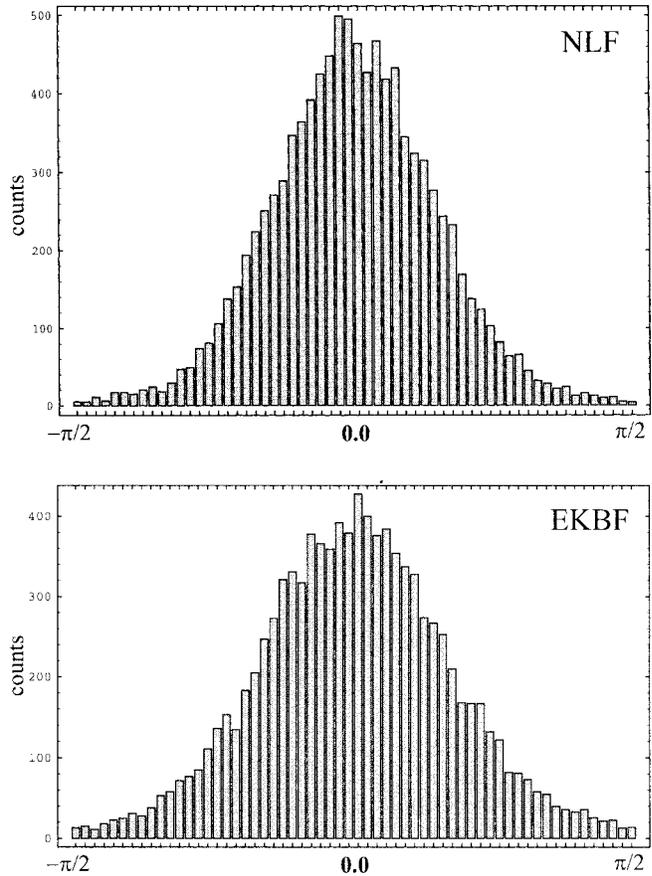


Fig. 9. Error histograms corresponding to the estimates of Figs. 7 (NLF) and 8 (EKBF).

the meaning of  $T_1$ ,  $T_2$ , and  $T_3$  being shown in Fig. 5, and the superscript  $T$  standing for vector transpose. The dimension of this state vector is

$$L \equiv (T_2 + T_3)T_1 - T_3 = T_1T_2 + (T_1 - 1)T_3. \quad (18)$$

The state vector  $\mathbf{v}_n$  is in fact all that is necessary to generate  $x_{i,j+1}$ , which is the first element of  $\mathbf{v}_{n+1}$ . Some other elements of  $\mathbf{v}_{n+1}$  are also in  $\mathbf{v}_n$ , so they can be obtained by index shifting operations. The  $T_1 - 1$  elements of  $\mathbf{v}_{n+1}$  which are not in  $\mathbf{v}_n$ , i.e., those in the rightmost column of the support  $\Theta_{l(n+1),c(n+1)}$ , have been previously generated and can be included as deterministic inputs. Formally, the reduced order state space model equivalent to (12), is then

$$\mathbf{v}_{n+1} = \mathbf{A}\mathbf{v}_n + \mathbf{B}w_n + \mathbf{E}\mathbf{e}_n \quad (19)$$

where  $\mathbf{A}$  is the  $L \times L$  state transition matrix containing the AR model parameters [the  $\theta_{i,j,k,l}$ 's in (12)],  $\mathbf{B}$  is the  $L \times 1$  matrix given by  $\mathbf{B} = [\mu \ 0 \ 0 \ \cdots \ 0]^T$ , the  $w_n$ 's are i.i.d. unit variance zero mean Gaussian variables (the *driving noise*),  $\mathbf{E}$  is a  $L \times (T_1 - 1)$  matrix with only zeros and ones, and

$$\mathbf{e}_n = \begin{bmatrix} x_{l(n)-1,c(n)+T_3+1} & x_{l(n)-2,c(n)+T_3+1} \\ \cdots & x_{l(n)-(T_1-1)c(n)+T_3+1} \end{bmatrix}^T \quad (20)$$

is the referred deterministic input.

The above-mentioned free boundary conditions of the AR model translate, in this state-space formalism, into arbitrarily-valued boundary pixels and a space variant state transition

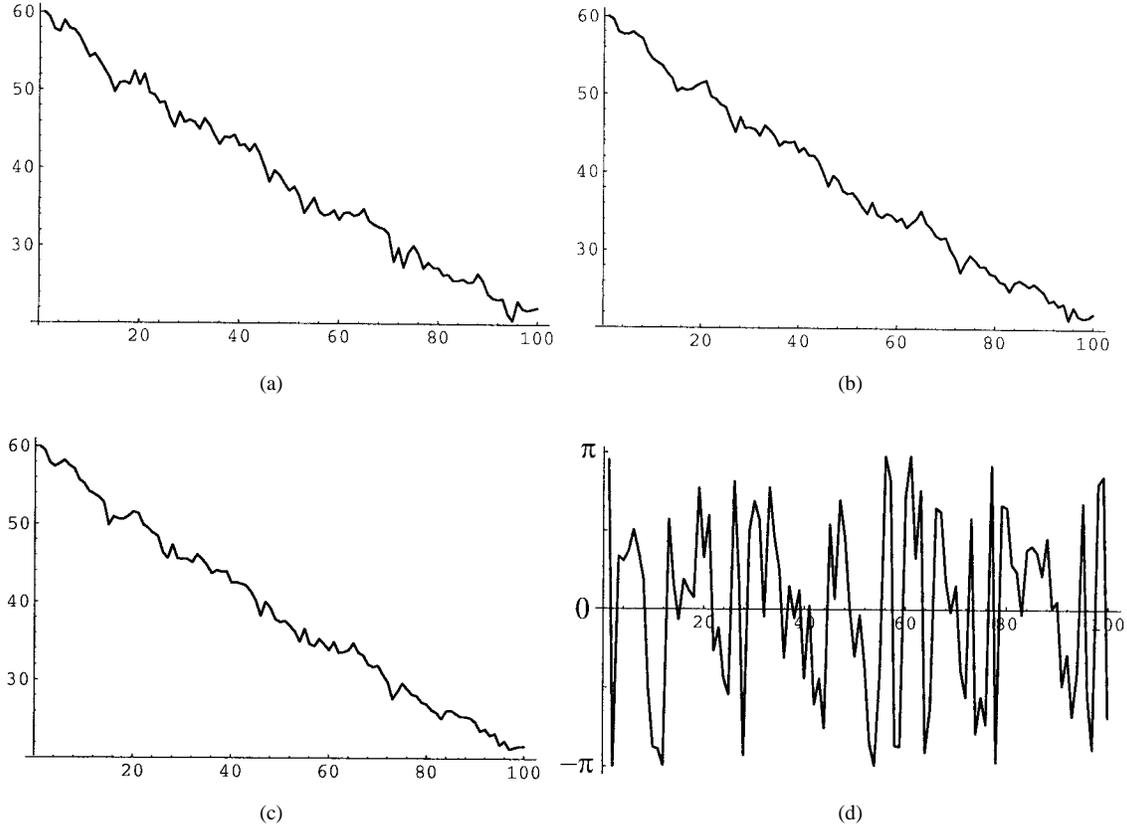


Fig. 10. Diagonal profiles. (a) Original surface (Fig. 2). (b) NLF estimate (Fig. 7). (c) EKBF estimate (Fig. 8). (d) Wrapped phases (the  $\eta_{i,j}$ 's shown in Fig. 4).

equation: when the coordinates  $(l(n), c(n))$  approach the image boundary, the state transition matrix  $\mathbf{A}$  and the deterministic input matrix  $\mathbf{E}$  are modified so that the (external) boundary pixels correspond to zero coefficients in these matrices. This space-variant behavior of the model only occurs near the image boundaries and so we will not clutter the notation in (19) by appending indices to matrices  $\mathbf{A}$  and  $\mathbf{E}$ . Of course, the implementation will have to take these aspects into account.

An example will help elucidate the notation adopted. Consider the following AR model for which  $T_1 = 2$ ,  $T_2 = 1$ , and  $T_3 = 2$  (yielding  $L = 4$ ):

$$x_{i,j} = \theta_1 x_{i,j-1} + \theta_2 x_{i-1,j-1} + \theta_3 x_{i-1,j} + \theta_4 x_{i-1,j+1} + u_{i,j}. \quad (21)$$

The equivalent reduced order state-space model is

$$\begin{aligned} \overbrace{\begin{bmatrix} x_{l(n),c(n)+1} \\ x_{l(n)-1,c(n)+1} \\ x_{l(n)-1,c(n)+2} \\ x_{l(n)-1,c(n)+3} \end{bmatrix}}^{\mathbf{v}_{n+1}} &= \overbrace{\begin{bmatrix} \theta_1 & \theta_2 & \theta_3 & \theta_4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}}^{\mathbf{A}} \overbrace{\begin{bmatrix} x_{l(n),c(n)} \\ x_{l(n)-1,c(n)} \\ x_{l(n)-1,c(n)+1} \\ x_{l(n)-1,c(n)+2} \end{bmatrix}}^{\mathbf{v}_n} \\ &+ \overbrace{\begin{bmatrix} \mu \\ 0 \\ 0 \\ 0 \end{bmatrix}}^{\mathbf{B}} w_n + \overbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}}^{\mathbf{E}} \overbrace{[x_{l(n)-1,c(n)+3}]}^{\mathbf{e}_n}. \end{aligned}$$

Concerning the boundary pixels, we adopt the following approach: in the first line, a 1-D AR model of the type  $x_{1,j} = \theta_5 x_{1,j-1} + u_{1,j}$  is used; for the first pixel of each

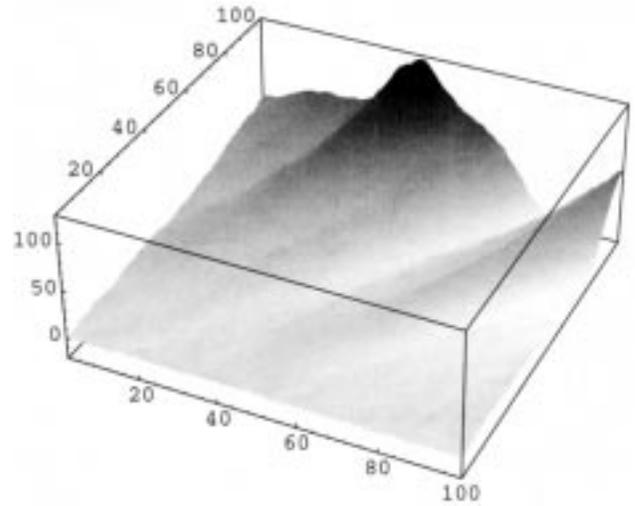


Fig. 11. Original phase surface/image (second example).

line the model is  $x_{i,1} = \theta_6 x_{i-1,1} + \theta_7 x_{i-1,2} + u_{i,1}$ ; the last pixel of each line is obtained according to  $x_{i,N} = \theta_8 x_{i-1,N-1} + \theta_9 x_{i-1,N} + u_{i,1}$ . The ROM matrices have to be adequately modified to reflect these boundary conditions.

Finally, notice that in terms of the state vector  $\mathbf{v}_n$ , the observation equation (6) becomes

$$\mathbf{y}_n \equiv \begin{bmatrix} y_{l(n),c(n)}^c \\ y_{l(n),c(n)}^s \end{bmatrix} = \begin{bmatrix} \cos(\mathbf{D}\mathbf{v}_n) \\ \sin(\mathbf{D}\mathbf{v}_n) \end{bmatrix} + \begin{bmatrix} n_{l(n),c(n)}^c \\ n_{l(n),c(n)}^s \end{bmatrix} \quad (22)$$

where  $\mathbf{D} = [1 \ 0 \ 0 \ \cdots \ 0]$ , so that  $\mathbf{D}\mathbf{v}_n = x_{l(n),c(n)}$ .

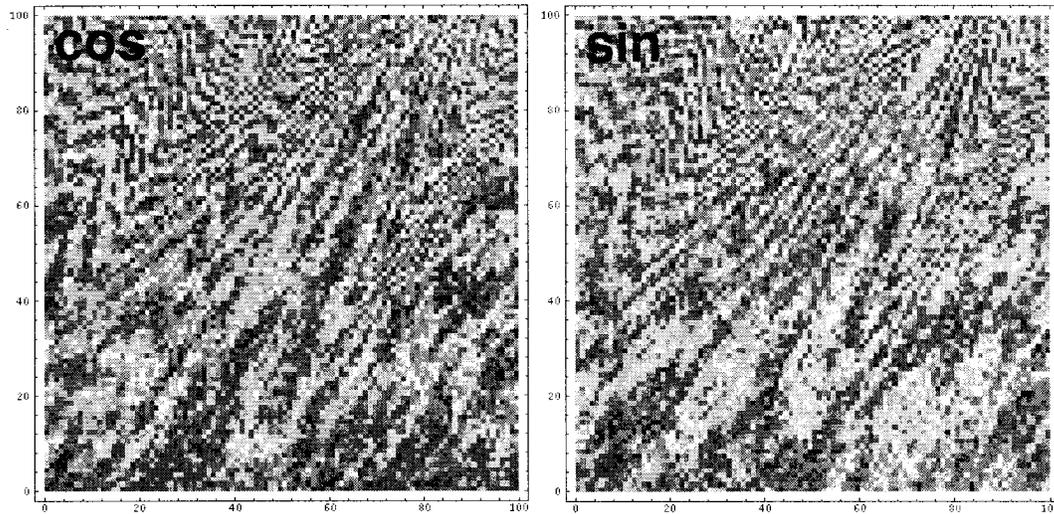


Fig. 12. Noisy ( $\sigma = 0.5$ ) cosine and sine observations of the original image/surface of Fig. 11.

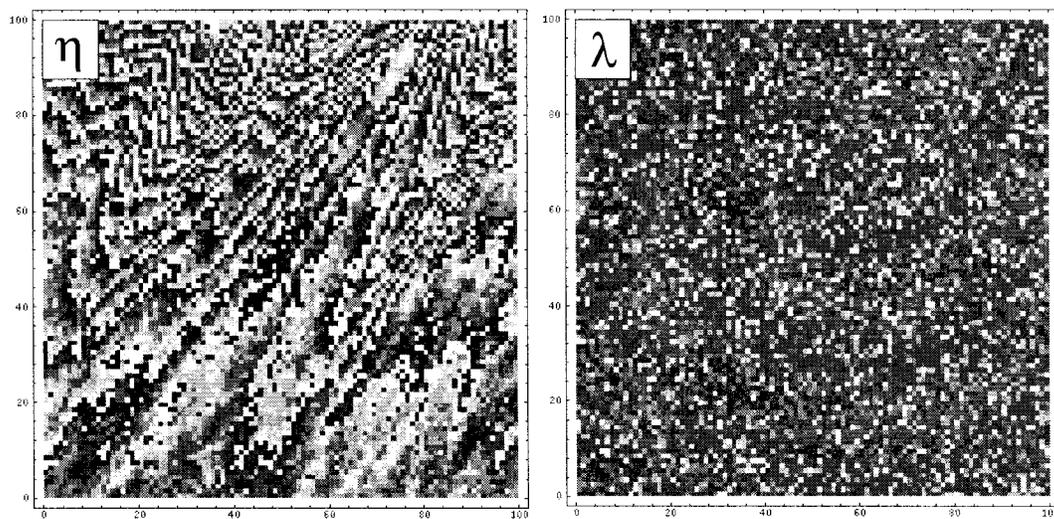


Fig. 13.  $\eta_{i,j}$ 's (interferogram) and  $\lambda_{i,j}$ 's obtained from the observations shown in Fig. 12.

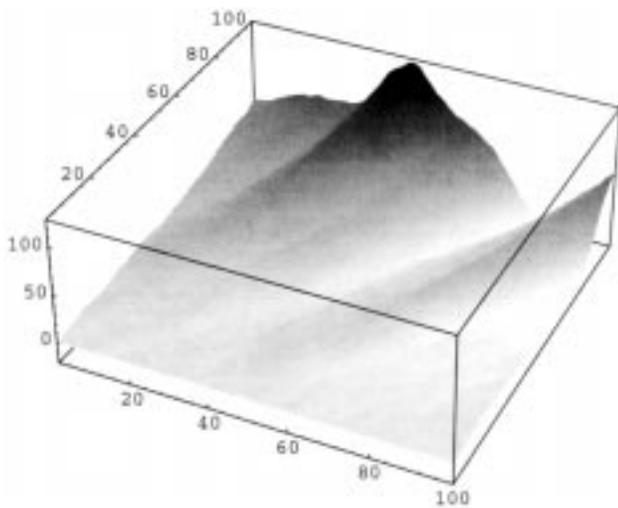


Fig. 14. Phase image/surface estimate obtained by the NLF from the noisy observations shown in Fig. 12 (compare with Fig. 11).

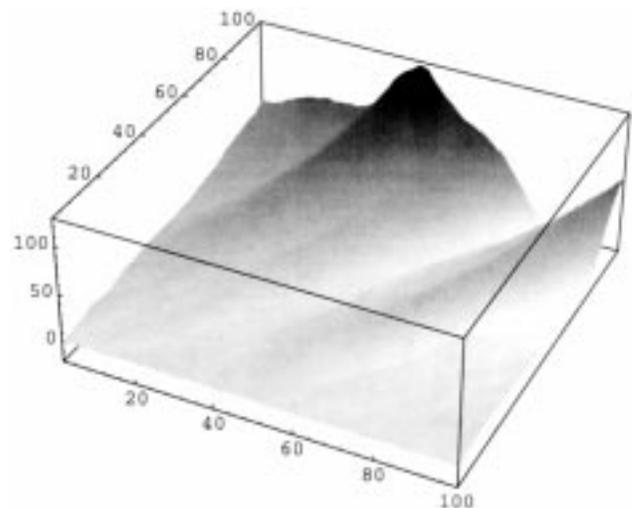


Fig. 15. Phase image/surface estimate obtained by the EKFB from the noisy observations shown in Fig. 12 (compare with Fig. 11).

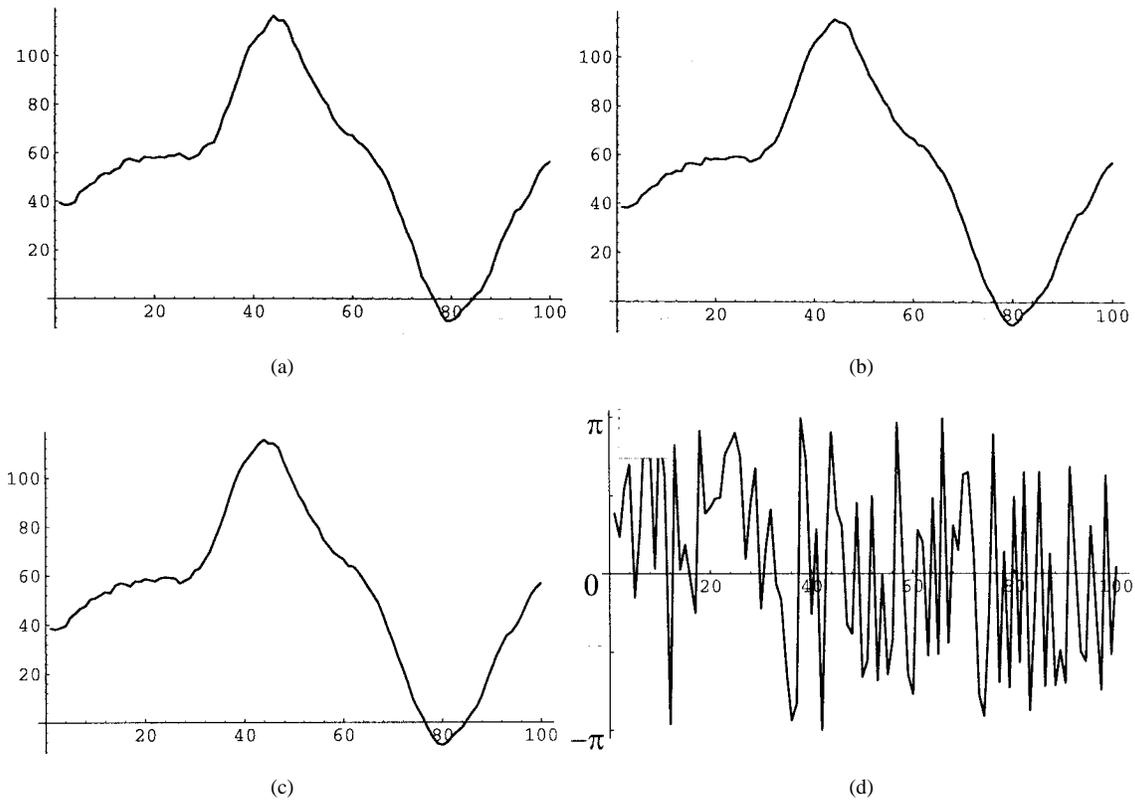


Fig. 16. Diagonal profiles. (a) Original surface (Fig. 11). (b) NLF estimate (Fig. 14). (c) EKBF estimate (Fig. 15). (d) Wrapped phases (the  $\eta_{i,j}$ 's shown in Fig. 13).

#### IV. PROPOSED SOLUTION

##### A. Introduction

Stochastic filtering is essentially a Bayesian approach to the recursive estimation a random process from its noisy observations. When the state transition (dynamics) and observation equations are linear, the driving and observation noises are Gaussian, and the initial condition is a Gaussian random variable, Kalman–Bucy (linear) filtering provides the optimal solution [22], [32]. When (at least) one of these conditions fails, the problem falls into the general framework of stochastic nonlinear filtering [7]–[9], [22], [32], [48].

##### B. General Optimal Solution

Given the models defined by (19) and (22), consider the problem of estimating  $\mathbf{v}_n$ , based on the set of “present” and “past” observations  $\mathbf{Y}_n = \{\mathbf{y}_k : k \leq n\}$ . In the Bayesian perspective, all the information concerning  $\mathbf{v}_n$  is contained in the conditional probability density function  $F_n = p(\mathbf{v}_n | \mathbf{Y}_n)$ , usually termed *filtering density*. A nonlinear filter propagates the filtering density by recursive (alternating) application of

$$\text{Prediction: } P_n = S_n * F_{n-1} \quad (23)$$

$$\text{Filtering: } F_n = C_n H_n \bullet P_n \quad (24)$$

where  $*$  denotes convolution and  $\bullet$  means pointwise multiplication (both defined on the state space); the arguments of

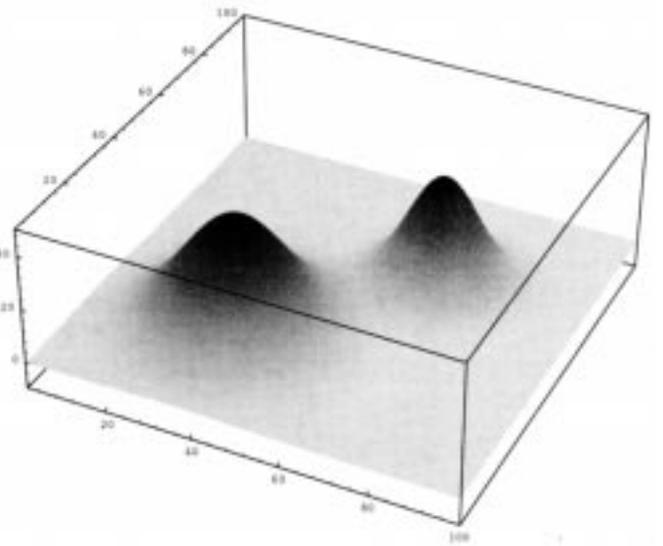


Fig. 17. Deterministic phase surface/image (third example).

the functions were omitted to simplify the notation. The other entities in (23)–(24) are as follows.

- The *convolution kernel*,  $S_n = p(\mathbf{v}_n | \mathbf{v}_{n-1})$  which, reflecting model (19), is Gaussian, i.e.,

$$S_n \propto \mathcal{N}(\mathbf{v}_n - \mathbf{A}\mathbf{v}_{n-1} - \mathbf{E}\mathbf{e}_{n-1}, \mathbf{B}\mathbf{B}^T) \quad (25)$$

where  $\mathcal{N}(\mathbf{s}, \mathbf{V}) \equiv \exp\{-\mathbf{s}^T \mathbf{V}^{-1} \mathbf{s} / 2\}$ ; recall that the  $w_n$ 's [see (19)] are assumed to have unit variance. The deterministic input  $\mathbf{e}_{n-1}$  contains previously obtained (in

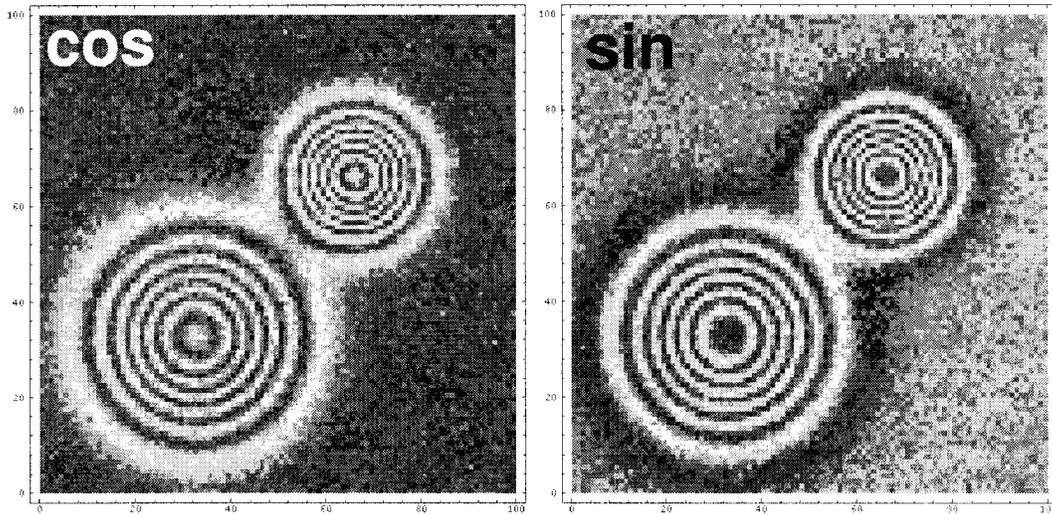


Fig. 18. Noisy ( $\sigma = 0.4$ ) cosine and sine observations of the original image/surface of Fig. 17.

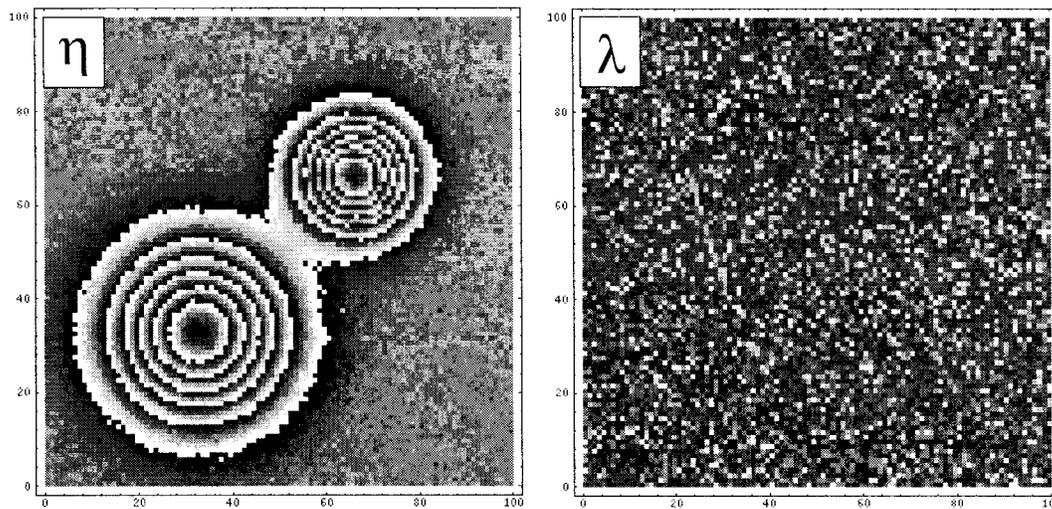


Fig. 19.  $\eta_{i,j}$ 's (interferogram) and  $\lambda_{i,j}$ 's obtained from the observations displayed in Fig. 18.

a lexicographical sense) estimates. Of course this is an approximation because it treats these field estimates as determinist, thus reducing their effect to a mean shift.

- The *observation (or sensor) factor*,  $H_n = p(\mathbf{y}_n | \mathbf{v}_n)$ , which, taking into account the observation model introduced in (8) in Section II, is given by

$$H_n \propto \exp\{\lambda_n \cos(\mathbf{D}\mathbf{v}_n - \eta_n)\} \quad (26)$$

where  $\eta_n \equiv \eta_{(n),c(n)}$  and  $\lambda_n \equiv \lambda_{l(n),c(n)}$ , with  $\eta_{i,j}$  and  $\lambda_{i,j}$  as given by (10) and (11).

- The *prediction density*,  $P_n = p(\mathbf{v}_n | \mathbf{Y}_{n-1})$ .
- A normalizing constant  $C_n$ .

For a detailed derivation of these equations see [7]–[9], [32], [48]. As a brief justification, insert (23) into (24) and explicitly rewrite the result as

$$H_n \bullet (S_n * F_{n-1}) \\ = p(\mathbf{y}_n | \mathbf{v}_n) p(\mathbf{v}_n | \mathbf{v}_{n-1}) * p(\mathbf{v}_{n-1} | \mathbf{Y}_{n-1}) \quad (27)$$

$$= p(\mathbf{y}_n | \mathbf{v}_n) \left( \int_{\mathbb{R}^L} p(\mathbf{v}_n | \mathbf{v}_{n-1}) p(\mathbf{v}_{n-1} | \mathbf{Y}_{n-1}) d\mathbf{v}_{n-1} \right) \quad (28)$$

$$= p(\mathbf{y}_n | \mathbf{v}_n) \left( \int_{\mathbb{R}^L} p(\mathbf{v}_n, \mathbf{v}_{n-1} | \mathbf{Y}_{n-1}) d\mathbf{v}_{n-1} \right) \quad (29)$$

$$= p(\mathbf{y}_n | \mathbf{v}_n, \mathbf{Y}_{n-1}) p(\mathbf{v}_n | \mathbf{Y}_{n-1}) \quad (30)$$

$$= p(\mathbf{v}_n | \mathbf{y}_n, \mathbf{Y}_{n-1}) p(\mathbf{y}_n | \mathbf{Y}_{n-1}) \quad (31)$$

$$\propto p(\mathbf{v}_n | \mathbf{Y}_n) = F_n. \quad (32)$$

Obtaining (29) from (28) requires noticing that  $p(\mathbf{v}_n, \mathbf{v}_{n-1} | \mathbf{Y}_{n-1}) = p(\mathbf{v}_n | \mathbf{v}_{n-1}, \mathbf{Y}_{n-1}) p(\mathbf{v}_{n-1} | \mathbf{Y}_{n-1})$  and that  $p(\mathbf{v}_n | \mathbf{v}_{n-1}, \mathbf{Y}_{n-1}) = p(\mathbf{v}_n | \mathbf{v}_{n-1})$  (i.e., with  $\mathbf{v}_{n-1}$  given,  $\mathbf{v}_n$  does not depend on  $\mathbf{Y}_{n-1}$ ). From (29) to (30), notice that  $p(\mathbf{y}_n | \mathbf{v}_n) = p(\mathbf{y}_n | \mathbf{v}_n, \mathbf{Y}_n)$  (i.e., conditioned on the current state, the current observation does not depend on the past ones). Equality between (30) and (31) is simply Bayes law. To write (32), notice that  $\mathbf{Y}_n = \{\mathbf{y}_n\} \cup \mathbf{Y}_{n-1}$  and that  $p(\mathbf{y}_n | \mathbf{Y}_{n-1})$  is a constant (with respect to  $\mathbf{v}_n$ ).

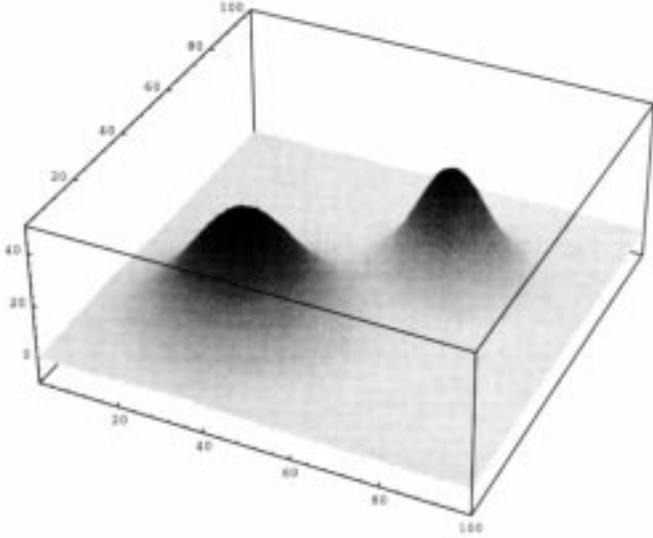


Fig. 20. Phase image/surface estimate obtained by the NLF from the noisy observations shown in Fig. 18 (compare with Fig. 17).

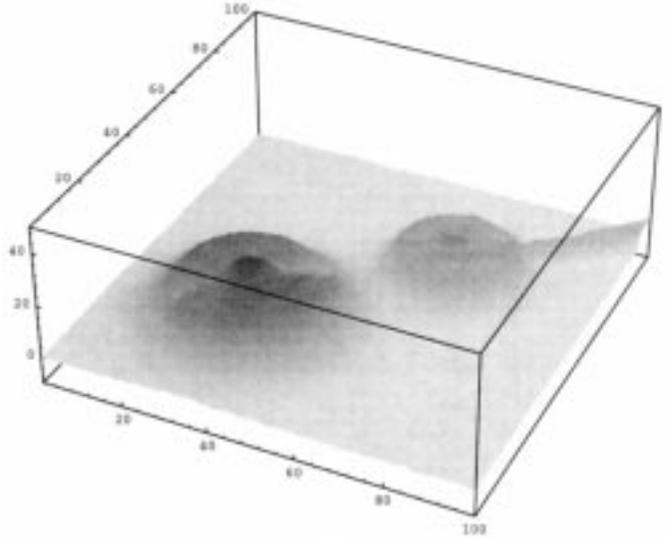


Fig. 21. Phase image/surface estimate obtained by the EKFB from the noisy observations shown in Fig. 18 (compare with Figs. 17 and 20).

### C. Implementation Aspects

To implement operations (23) and (24), approximate finite representations of the operands are required. Each (normalized) period of  $H_n$  is a Tikhonov function [56] that tends to a Gaussian function for large values of  $\lambda_n$  and becomes flat as  $\lambda_n$  goes to zero. This fact suggests representing the periodic positive function  $H_n$  by a train of Gaussian terms

$$\tilde{H}_n \propto \sum_{i=-\infty}^{+\infty} \mathcal{N}(\mathbf{D}\mathbf{v}_n - \eta_n^{(i)}, \gamma^{\tilde{H}_n}) \quad (33)$$

centered on  $\eta_n^{(i)} \equiv \eta_n + 2\pi i$  (the maxima of  $H_n$ ) with common variance  $\gamma^{\tilde{H}_n}$  (notice that  $\mathbf{D}\mathbf{v}_n$  is scalar) [40], [41]. Representation (33) should reproduce, as much as possible, the shape of  $H_n$ , for all values of  $\lambda_n$ . As in [40] and [41], we adopt the following minimum Kullback divergence criterion.

- Take  $\eta_n^{(0)} = 0$  for simplicity, and consider the normalized central periods of  $H_n$  and  $\tilde{H}_n$ . Denote these probability density functions as  $h(x, \lambda)$  and  $\tilde{h}(x, \gamma)$ , respectively, where the recursion index  $n$  dropped and  $\gamma^{\tilde{H}_n}$  is simply referred to as  $\gamma$ .
- Consider the Kullback divergence

$$D(h(x, \lambda) \parallel \tilde{h}(x, \gamma)) \equiv \int_{-\pi}^{\pi} h(x, \lambda) \log \frac{h(x, \lambda)}{\tilde{h}(x, \gamma)} dx, \quad (34)$$

a dissimilarity measure between the probability density functions  $h(x, \lambda)$  and  $\tilde{h}(x, \gamma)$  [35], [40].

- For any given  $\lambda$ , obtain the optimal representation  $\tilde{h}(x, \gamma)$  by determining the value of  $\gamma$  that minimizes  $D(h(x, \lambda) \parallel \tilde{h}(x, \gamma))$  [40].
- Since the Kullback divergence is a convex functional with respect to both probability density functions involved [35], the minimum can be found by computing its partial derivative with respect to  $\gamma$  and setting it equal to zero;

this leads to

$$\int_{-\pi}^{\pi} h(x, \lambda) \frac{\partial \tilde{h}(x, \gamma)}{\partial \gamma} \frac{1}{\tilde{h}(x, \gamma)} dx = 0. \quad (35)$$

- Inserting the particular forms of  $h(x, \lambda)$  and  $\tilde{h}(x, \gamma)$  into (35), yields a nonlinear integral equation. A look-up table of numerical solutions, built off-line, can be accessed with a minor computational effort.

Notice that only two parameters are needed to represent  $H_n$ : the location of a maximum of  $H_n$  [given by (11)] and the common variance  $\gamma^{\tilde{H}_n}$  obtained from this look-up table. To have an idea of the quality of these representations see the related figures in [40].

At each step  $n$ , the estimate  $\hat{\mathbf{v}}_n$  is obtained by minimizing the expectation (with respect to the filtering density  $F_n$ ) of an adopted cost function  $C(\mathbf{v}_n - \hat{\mathbf{v}}_n)$  [7], [32].

### D. Prediction

Assume that  $F_{n-1}$  is a finite sum of  $N^{F_{n-1}}$  Gaussian functions

$$\begin{aligned} F_{n-1} &= p(\mathbf{v}_{n-1} \mid \mathbf{Y}_{n-1}) \\ &= \sum_{i=1}^{N^{F_{n-1}}} k_i^{F_{n-1}} \mathcal{N}(\mathbf{v}_{n-1} - \boldsymbol{\eta}_i^{F_{n-1}}, \mathbf{V}^{F_{n-1}}) \end{aligned} \quad (36)$$

with means  $\boldsymbol{\eta}_i^{F_{n-1}} \in \mathbb{R}^L$ , common covariance matrix  $\mathbf{V}^{F_{n-1}}$  (dimension  $L \times L$ ), and the weighting factors  $k_i^{F_{n-1}}$  such that  $F_{n-1}$  is a normalized density. According to (23) [see also (28)–(30)], the prediction density  $P_n = p(\mathbf{v}_n \mid \mathbf{Y}_{n-1})$  is obtained by the convolution of  $F_{n-1}$  [given by (36)] with the kernel  $S_n = p(\mathbf{v}_n \mid \mathbf{v}_{n-1})$  [given by (25)], i.e., by computing the integral inside the parenthesis in (28). The result is a finite sum of  $N^{P_n}$  Gaussian terms

$$P_n = \sum_{i=1}^{N^{P_n}} k_i^{P_n} \mathcal{N}(\mathbf{v}_n - \boldsymbol{\eta}_i^{P_n}, \mathbf{V}^{P_n}) \quad (37)$$

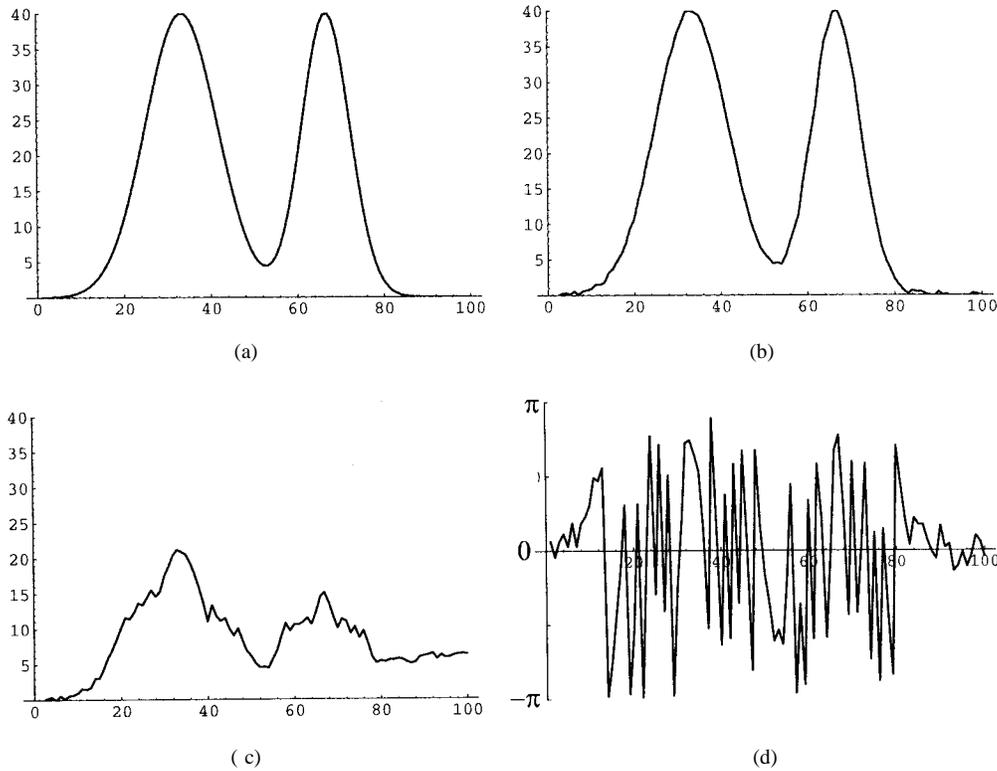


Fig. 22. Diagonal profiles. (a) Original surface of Fig. 17. (b) Estimate of Fig. 20 (nonlinear algorithm). (c) Estimate of Fig. 21 (EKBF algorithm). (d) Wrapped phases (the  $\eta_{i,j}$ 's) of Fig. 19.

where

$$\boldsymbol{\eta}_i^{P_n} = \mathbf{A}\boldsymbol{\eta}_i^{F_{n-1}} + \mathbf{E}\mathbf{e}_{n-1} \quad (38)$$

$$\mathbf{V}^{P_n} = \mathbf{A}\mathbf{V}^{F_{n-1}}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T \quad (39)$$

$$k_i^{P_n} = Ck_i^{F_{n-1}}. \quad (40)$$

Constant  $C$  affecting all the terms of  $P_n$  can be omitted. Since (25) is a single Gaussian, there is no creation of new modes, i.e.,  $N^{P_n} = N^{F_{n-1}}$ . Prediction is thus a bank of  $N^{F_{n-1}}$  discrete Kalman–Bucy prediction steps [32].

### E. Filtering

To implement the filtering step (24), the representation  $\tilde{H}_n$ , given by (33), is used instead of  $H_n$ . Since multiplying two Gauss functions yields a third Gauss function, the result of multiplying the  $i$ th term of  $P_n$  by the  $l$ th mode of  $\tilde{H}_n$ , denoted by  $(P_n \bullet \tilde{H}_n)_{i,l}$ , is itself Gaussian; specifically

$$(P_n \bullet \tilde{H}_n)_{i,l} \propto k_{i,l}^{F_n} \mathcal{N}(\mathbf{v}_n - \boldsymbol{\eta}_{i,l}^{F_n}, \mathbf{V}^{F_n}) \quad (41)$$

with

$$\boldsymbol{\eta}_{i,l}^{F_n} = \boldsymbol{\eta}_i^{P_n} - \frac{\mathbf{D}\boldsymbol{\eta}_i^{P_n} - \boldsymbol{\eta}_n^{(l)}}{V_{1,1}^{P_n} + \gamma\tilde{H}_n} \begin{bmatrix} V_{1,1}^{P_n} \\ \vdots \\ V_{1,L}^{P_n} \end{bmatrix} \quad (42)$$

$$\mathbf{V}^{F_n} = \mathbf{V}^{P_n} - \frac{1}{V_{1,1}^{P_n} + \gamma\tilde{H}_n} \begin{bmatrix} V_{1,1}^{P_n} & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V_{1,L}^{P_n} & 0 & \dots & 0 \end{bmatrix} \mathbf{V}^{P_n} \quad (43)$$

$$k_{i,l}^{F_n} = k_i^{P_n} \mathcal{N}(\mathbf{D}\boldsymbol{\eta}_i^{P_n} - \boldsymbol{\eta}_n^{(l)}, V_{1,1}^{P_n} + \gamma\tilde{H}_n) \quad (44)$$

where  $\mathbf{D}\boldsymbol{\eta}_i^{P_n}$  is the first component of  $\boldsymbol{\eta}_i^{P_n}$ , and  $V_{m,n}$  stands for the element  $(m,n)$  of matrix  $\mathbf{V}$ .

Equations (42) and (43) correspond, formally, to an infinite bank of Kalman–Bucy filtering steps [weighted according to (44)], each one guided by the (pseudo)observations  $\eta_n^{(i)}$  with the same data-dependent (equivalent) “noise variance”  $\gamma\tilde{H}_n$ . Notice that the covariance matrix  $\mathbf{V}^{F_n}$  in (43) is not subscripted by the index  $i$ , i.e., all the modes of  $F_n$  have the same covariance matrix.

Straight application of the filtering step would produce an infinite number of Gaussian terms; in the implementations described in [41] and [49], the dimension of the filter (i.e., the number of Gaussian terms) is controlled by the following mechanisms.

- Each mode  $i$  of  $P_n$  is multiplied by only the  $J$  nearest modes of  $\tilde{H}_n$ ; i.e., in (42), (43), and (44),  $l$  ranges over only  $J$  different values. Call  $J$  the *multiplication parameter*.
- In the resulting function, those modes which are closer than some threshold distance are agglutinated, while those which weight less than another threshold are eliminated (pruned).

Adjusting the corresponding control parameters is an important issue which depends on many aspects of the particular model considered. Another critical point is the choice of an estimation criterion (cost function) considering the multimodal shape of the filtering density; [41] and [49] adopted the minimum mean squared error criterion which, in some cases, may not lead to good *local* (modulo- $2\pi$ ) estimates.

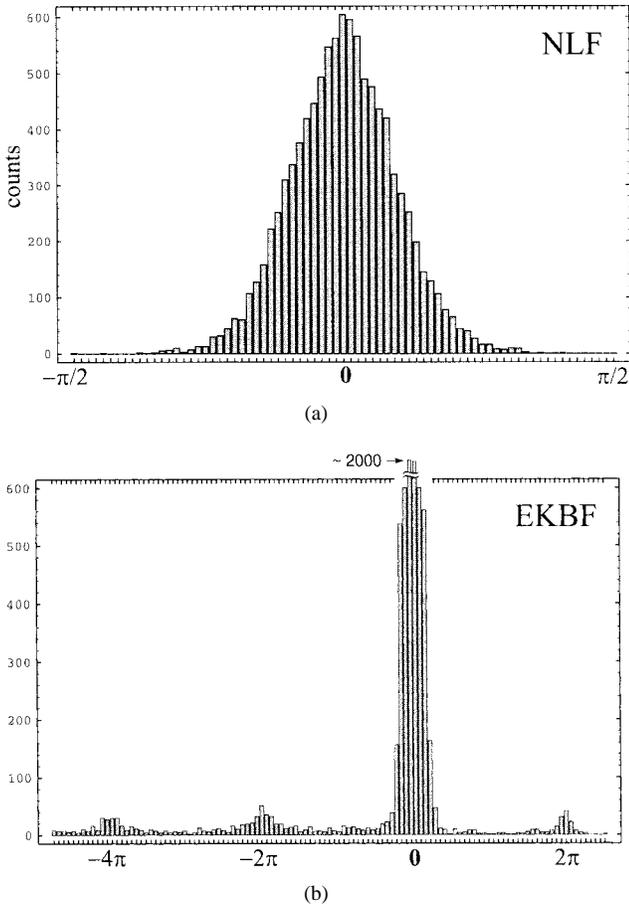


Fig. 23. Error histograms corresponding to the estimates of Figs. 20 and 21; notice the multimodal nature of the second histogram due to the  $2\pi$  phase jumps.

### F. Simplification

To avoid the above details and the corresponding implementation complexity, we adopt, in this paper, the simplification next described. The key feature lies in setting the multiplication parameter equal to one,  $J = 1$ . Assuming that the prediction density  $P_n$  is unimodal (i.e.,  $N^{P_n} = 1$ ), the filtering density, which is obtained by multiplying  $P_n$  only by the nearest mode (since  $J = 1$ ) of  $\tilde{H}_n$ , will also be a single Gaussian. Agglutination and pruning loose any sense and the next prediction density  $P_{n+1}$  will automatically be unimodal. Furthermore, the Gaussianity of this filtering density implies that the optimal estimate, with respect to any reasonable cost function, is simply its mean.

The complete final structure of the (simplified) NLF is next described. As just stated, assume that  $P_n$  is a single Gaussian ( $N^{P_n} = 1$ ) with mean  $\eta^{P_n}$  and covariance matrix  $\mathbf{V}^{P_n}$ .

*NLF Algorithm:*

- Step 1:** Given  $\lambda_n = \lambda_{l(n),c(n)}$ , function of the observed data via (10), consult the look-up table (see Section IV-C) to find the optimal (in a minimum Kullback divergence sense)  $\gamma^{\tilde{H}_n}$ .
- Step 2:** Find which mode of the representation of the observation factor (i.e., of  $\tilde{H}_n$ ) is closer to the first component of the prediction density (see Fig. 6).

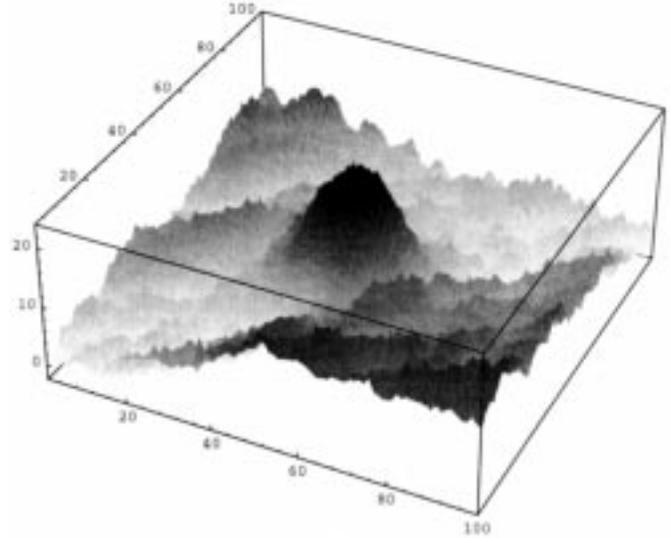


Fig. 24. Surface obtained by superimposing a Gaussian “hill” with a sample of model (54).

$$\begin{aligned} \text{Call it } \eta_n^{(l^*)} &= \eta_n + 2\pi l^*, \text{ where } l^* \text{ is given by} \\ l^* &= \arg \min_l \{ (\mathbf{D}\eta^{P_n} - (\eta_n + 2l\pi))^2 \} \\ &= \text{round} \left[ \frac{\mathbf{D}\eta^{P_n} - \eta_n}{2\pi} \right]. \end{aligned} \quad (45)$$

**Step 3:** Correct  $\eta_n^{(l^*)}$  according to

$$\eta^{F_n} = \eta^{P_n} - \frac{\mathbf{D}\eta^{P_n} - \eta_n^{(l^*)}}{V_{1,1}^{P_n} + \gamma^{\tilde{H}_n}} \begin{bmatrix} V_{1,1}^{P_n} \\ \vdots \\ V_{1,L}^{P_n} \end{bmatrix} \quad (46)$$

which is (42) without the indices  $i$  (since  $N^{P_n} = 1$ ) and  $l$  (because  $J = 1$ ). This corrected value is the estimate  $\hat{\nu}_n = \eta^{F_n}$ .

**Step 4:** Compute the covariance matrix  $\mathbf{V}^{F_n}$  according to (43).

**Step 5:** Compute the parameters of the (Gaussian) prediction density, as follows.

- Mean:  $\eta^{P_{n+1}} = \mathbf{A}\eta^{F_n} + \mathbf{E}\mathbf{e}_n$ , according to (38).
- Covariance:  $\mathbf{V}^{P_{n+1}} = \mathbf{A}\mathbf{V}^{F_n}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T$ , according to (39)

where  $\mathbf{e}_n$  contains previously obtained estimates, according to its definition (20).

**Step 6:** Increment  $n$  by one and go back to **Step 1**.

Notice that Steps 1, 2, 3, and 4 implement the filtering operation (24), while Step 5 implements the prediction (23). Steps 2 and 3 are illustrated in Fig. 6.

### G. Extended Kalman–Bucy Filter

Assume a Gaussian prediction density, now denoted as  $\bar{P}_n$ , with mean  $\eta^{\bar{P}_n}$  and covariance  $\mathbf{V}^{\bar{P}_{n+1}}$ . The extended EKBF is obtained by linearizing the nonlinear observation model (22) around  $\eta^{\bar{P}_n}$ . As a consequence of this linearization, the sensor factor, now denoted as  $\bar{H}_n$ , loses its multimodal structure

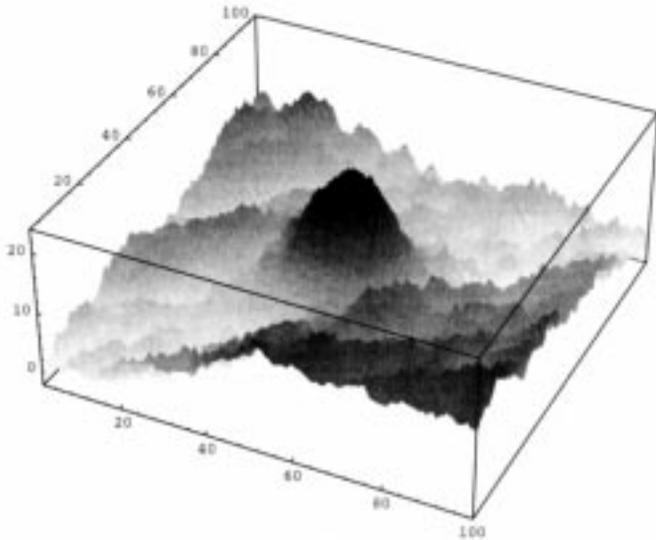


Fig. 25. Reconstruction of the surface of Fig. 24, obtained by the NLF.

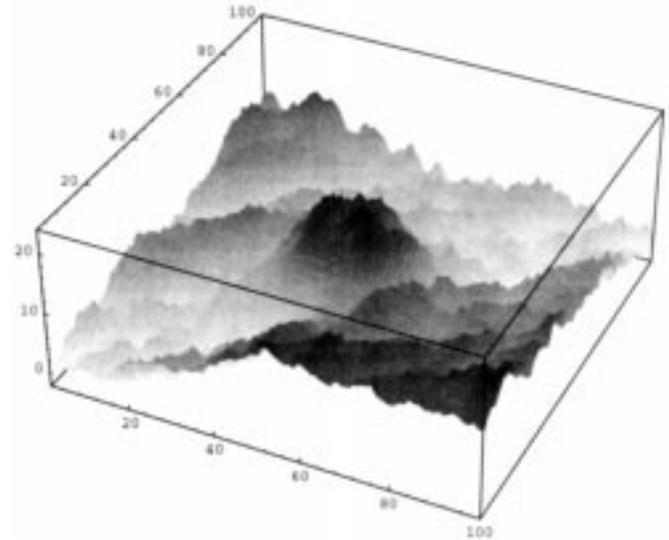


Fig. 26. Reconstruction of the surface of Fig. 24, obtained by the EKBF.

reducing to the Gaussian form

$$\tilde{H}_n \propto \mathcal{N}(\mathbf{D}\mathbf{v}_n - \eta^{\tilde{H}_n}, \sigma_n^2) \quad (47)$$

where  $\sigma_n^2 \equiv \sigma_{l(n),c(n)}^2$  is the observation noise variance at pixel  $(l(n), c(n))$ , and

$$\eta^{\tilde{H}_n} = \mathbf{D}\eta^{\tilde{F}_n} + y_n^s \cos(\mathbf{D}\eta^{\tilde{F}_n}) - y_n^c \sin(\mathbf{D}\eta^{\tilde{F}_n}). \quad (48)$$

The filtering density, now  $\tilde{F}_n$ , is obviously also a single Gaussian. The structure of the EKBF is next described.

*EKBF Algorithm:*

**Step 1: Filtering.** Compute the mean and covariance of the filtering density according to

$$\eta^{\tilde{F}_n} = \eta^{\tilde{F}_n} + \frac{y_n^s \cos(\mathbf{D}\eta^{\tilde{F}_n}) - y_n^c \sin(\mathbf{D}\eta^{\tilde{F}_n})}{V_{1,1}^{\tilde{F}_n} + \sigma_n^2} \begin{bmatrix} V_{1,1}^{\tilde{F}_n} \\ \vdots \\ V_{1,L}^{\tilde{F}_n} \end{bmatrix} \quad (49)$$

$$\mathbf{V}^{\tilde{F}_n} = \mathbf{V}^{\tilde{F}_n} - \frac{1}{V_{1,1}^{\tilde{F}_n} + \sigma_n^2} \begin{bmatrix} V_{11}^{\tilde{F}_n} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ V_{1,L}^{\tilde{F}_n} & 0 & \cdots & 0 \end{bmatrix} \mathbf{V}^{\tilde{F}_n}. \quad (50)$$

The state estimate at each step is, simply,  $\hat{\mathbf{v}}_n = \eta^{\tilde{F}_n}$ .

**Step 2: Prediction.** Convolving the Gaussian filtering density  $\tilde{F}_n$  with (25), one has again a Gaussian prediction density,  $\tilde{F}_{n+1}$ , with mean and covariance matrix given, respectively, by

$$\eta^{\tilde{F}_{n+1}} = \mathbf{A}\eta^{\tilde{F}_n} + \mathbf{E}\mathbf{e}_n \quad (51)$$

$$\mathbf{V}^{\tilde{F}_{n+1}} = \mathbf{A}\mathbf{V}^{\tilde{F}_n}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T. \quad (52)$$

**Step 3:** Increment  $n$  by one and go back to **Step 1**.

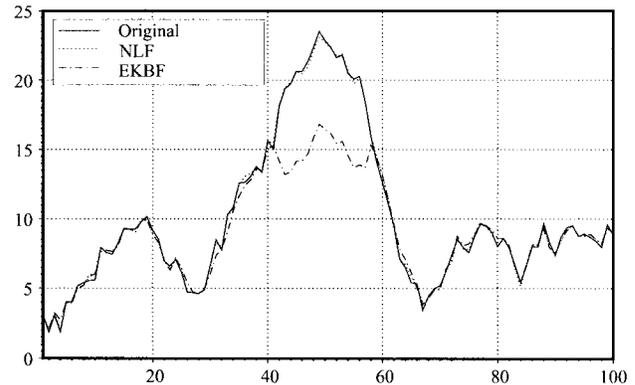


Fig. 27. Profiles of the surfaces of Figs. 24–26.

## V. EXPERIMENTAL RESULTS

The set of simulations next presented are not extensive nor statistically meaningful studies. Examples have been selected as they correspond to plausible real world situations and/or illustrate special features of the approach and algorithms here described.

The first experimental test considers the original image/surface of Fig. 2, which was generated according to (14). The noisy (with  $\sigma = 0.5$ ) real (cosine) and imaginary (sine) observations are presented in Fig. 3, while Fig. 4 displays the  $\eta_{i,j}$ 's (i.e., the wrapped phase values, or interferogram) and the  $\lambda_{i,j}$ 's. The estimate produced by the NLF is presented in Fig. 7, while Fig. 8 shows the estimate obtained by the EKBF, both visually indistinguishable from the original. However, the errors between the original image and the estimates, whose approximately Gaussian shaped histograms are presented in Fig. 9, have different standard deviations: 0.485 for the NLF, and 0.581 for the EKBF. Fig. 10 shows diagonal profiles of the original and estimated surfaces and also of the wrapped phase values.

The second test regards an unstable AR model,

$$x_{i,j} = 0.51x_{i,j-1} + 0.21x_{i-1,j} + 0.31x_{i-1,j-1} + u_{i,j} \quad (53)$$

with the standard deviation of the driving noise set to  $\mu = 0.75$ ; a sample surface is shown in Fig. 11. The noisy observed images ( $\sigma = 0.5$ ) are in Fig. 12; Fig. 13 shows the  $\eta_{i,j}$ 's and the  $\lambda_{i,j}$ 's. The surfaces estimated by the NLF and the EKBF algorithms, displayed in Figs. 14 and 15, respectively, are again visually indistinguishable. The histograms of the errors (not shown) are again approximately Gaussian with different standard deviations: 0.529 for the NLF and 0.621 for the EKBF. In Fig. 16, profiles obtained along line 98 of the surfaces are shown. Notice how, despite the unstable nature of the phase surface model, both algorithms are able to follow its absolute value. In this example, there are phase differences between adjacent pixels exceeding  $2\pi$ ; many phase unwrapping methods often adopt the criterion that adjacent phase values can not differ by more than  $\pi$ . This surface could never be recovered by such techniques.

The next example illustrates the ability of the proposed approach to reconstruct surfaces for which there is no knowledge about the generating model. The surface to be reconstructed was deterministically generated as the sum of two 2-D Gauss functions, as shown in Fig. 17. The observations, with  $\sigma = 0.3$ , are shown in Fig. 18; the  $\eta_{i,j}$ 's and the  $\lambda_{i,j}$ 's are displayed in Fig. 19. The algorithms were implemented assuming the following *a priori* model:

$$x_{i,j} = 0.5x_{i,j-1} + 0.5x_{i-1,j} + u_{i,j} \quad (54)$$

with unit variance driving noise. This simple model is enough to express, in a formal way, the desired phase surface “continuity”; i.e., that each pixel is expected to be close to the average of its two nearest causal neighbors. The surface estimates obtained with the NLF and the EKBF are exhibited in Figs. 20 and 21, respectively. In this example, only the NLF succeeds in estimating the absolute phase surface, as is evident from the profiles shown in Fig. 22 and the error histograms presented in Fig. 23.

In the last example, a mixed situation is studied (similar to the one in [39]); a sample of a random surface generated according to (54) (with driving noise standard deviation  $\mu = 0.7$ ) is superimposed on a deterministic Gaussian shaped “hill” (see Fig. 24). Both filters were implemented assuming model (54); as in the previous example, only the NLF was able to adequately estimate the original surface. This is clear from the reconstructions shown in Figs. 25 and 26, and from the profile displayed in Fig. 27.

## VI. FINAL REMARKS

This paper formulated absolute phase imaging as a Bayesian nonlinear image reconstruction problem. The considered observation model captures the essential common features of a wide range of phase imaging techniques, appearing namely in InSAR, MRI, optical interferometry, and diffraction tomography. The prior model adopted for the original image is a nonsymmetrical half plane autoregressive (NSHP AR) Gauss Markov random field (GMRF). A reduced order state space formulation of this model allowed casting the problem into a stochastic nonlinear filtering framework. Specifically, nonlinear absolute phase estimation algorithms (studied in

[40], [41], and [49]) were adapted to this 2-D problem. The extended Kalman–Bucy filter (EKBF), a classical approach to stochastic nonlinear filtering, was also applied to the problem under study.

In the four examples presented, the nonlinear filter was able to adequately estimate absolute phase images; the EKBF failed in two of the examples and presented about 20% higher error standard deviation in the other two. This behavior was expectable from what has been shown in systematic studies on 1-D phase estimation problems (see [41] and references therein): nonlinear filters, designed to capture essential information contained in the conditional probability density functions, perform better than the EKBF, where linearization of the observation model may destroy important information.

In this paper, we have not addressed the setting of the involved model parameters, which is a critical issue of any model-based technique. One approach to this problem, which we intend to explore in future work, is to embed the proposed algorithm as part of an adaptive scheme, such as those in, e.g., [18], [19], and [37].

## REFERENCES

- [1] C. Allen, “Interferometric synthetic aperture radar,” *IEEE Geosci. Remote Sensing Soc. Newslett.*, pp. 6–13, Sept. 1995.
- [2] D. Angwin and H. Kaufman, “Image restoration using reduced order models,” *Signal Process.*, vol. 16, pp. 21–28, 1989.
- [3] ———, “Nonhomogeneous image identification and restoration procedures,” in *Digital Image Restoration*, A. Katsaggelos, Ed. New York: Springer-Verlag, 1991, pp. 177–208.
- [4] L. Axel and D. Morton, “Correction of phase wrapping in magnetic resonance imaging,” *Med. Phys.*, vol. 16, pp. 284–287, 1989.
- [5] N. Balram and J. Moura, “Recursive enhancement of noncausal images,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1991, pp. 2997–3000.
- [6] J. Besag, “On the statistical analysis of dirty pictures,” *J. R. Stat. Soc. B*, vol. 48, pp. 259–302, 1986.
- [7] R. Bucy, *Lectures on Discrete Time Filtering*. New York: Springer-Verlag, 1994.
- [8] R. Bucy, C. Hecht, and K. D. Senne, “An engineer’s guide to building nonlinear filters,” Tech. Rep. SRL-TR-72-0004, Frank J. Seiler Res. Lab., vols. 1–2, 1972.
- [9] R. Bucy and P. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, 2nd ed. New York: Chelsea, 1987.
- [10] R. Chellappa, “Two-dimensional discrete Gaussian Markov random field models for image processing,” in *Progress in Pattern Recognition*, L. Kanal and A. Rosenfeld, Eds. New York: Elsevier, 1985.
- [11] R. Chellappa, T. Simchony, and Z. Lichtenstein, “Image estimation using 2D noncausal Gauss–Markov random field models,” in *Digital Image Restoration*, A. Katsaggelos, Ed. New York: Springer-Verlag, 1991, pp. 109–140.
- [12] N. Ching, D. Rosenfeld, and M. Braun, “Two-dimensional phase unwrapping using a minimum spanning tree,” *IEEE Trans. Image Processing*, vol. 1, pp. 355–365, 1992.
- [13] H. Derin and P. Kelly, “Discrete-index Markov-type random processes,” in *Proc. IEEE*, vol. 77, pp. 1485–1510, 1989.
- [14] A. Devaney, “Geophysical diffraction tomography,” *IEEE Trans. Geosci. Remote Sensing*, vol. GE-22, pp. 3–13, 1984.
- [15] ———, “Diffraction tomographic reconstruction from intensity data,” *IEEE Trans. Image Processing*, vol. 1, pp. 221–228, 1992.
- [16] R. Dubes and A. Jain, “Random field models for image analysis,” *J. Appl. Stat.*, vol. 6, pp. 131–164, 1989.
- [17] S. Dupont and M. Berthod, “Interférométrie radar et déroulement de phase,” Tech. Rep. 2344, Inst. Nat. Rech. Inform. Automat., Sophia Antipolis, France, 1994.
- [18] M. Figueiredo and J. Leitão, “Adaptive discontinuity location in image restoration,” in *Proc. First IEEE Int. Conf. Image Processing*, Austin, TX, 1994, vol. II, pp. 665–669.
- [19] ———, “Unsupervised image restoration and edge location using compound Gauss–Markov random fields and the MDL principle,” *IEEE Trans. Image Processing*, vol. 6, pp. 1089–1102, 1997.

- [20] B. Friedlander and J. Francos, "Model based phase unwrapping of 2D signals," *IEEE Trans. Signal Processing*, vol. 44, pp. 2999–3007, 1996.
- [21] A. Gabriel and R. Goldstein, "Crossed orbit interferometry: Theory and experimental results from sir-b," *Int. J. Remote Sensing*, vol. 9, pp. 857–872, 1988.
- [22] A. Gelb, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.
- [23] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, pp. 721–741, 1984.
- [24] R. Goldstein and H. Zebker, "Interferometric radar measurements of ocean surface currents," *Nature*, vol. 328, pp. 707–709, 1987.
- [25] R. Goldstein, H. Zebker, and C. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, pp. 713–720, 1988.
- [26] L. Graham, "Synthetic interferometer radar for topographic mapping," *Proc. IEEE*, vol. 62, pp. 763–768, 1974.
- [27] A. Gray and P. Farris-Manning, "Repeat-pass interferometry with airborne synthetic aperture radar," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 180–191, 1993.
- [28] J. Gu, Y. Hung, and F. Chen, "Iteration algorithm for computer-aided speckle interferometry," *Appl. Opt.*, vol. 33, pp. 5308–5317, 1994.
- [29] C. Guarino, "Weighted two-dimensional phase unwrapping," in *Proc. 1995 Int. Geoscience and Remote Sensing Symp.*, 1995, pp. 193–195.
- [30] M. Hedley and D. Rosenfeld, "A new two-dimensional phase unwrapping algorithm for MRI images," *Magnet. Reson. Med.*, vol. 24, pp. 177–181, 1992.
- [31] K. Ho and J. Kahn, "Exact probability density function for phase measurement interferometry," *J. Opt. Soc. Amer. A*, vol. 12, pp. 1984–1989, 1995.
- [32] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
- [33] H. Kaufman and M. Tekalp, "Survey of estimation techniques in image restoration," *IEEE Control Syst. Mag.*, pp. 16–24, Jan. 1991.
- [34] M. Kaveh and M. Soumekh, "Computer-assisted diffraction tomography," in *Image Recovery: Theory and Application*, H. Stark, Ed. New York: Academic, pp. 369–413, 1987.
- [35] S. Kullback, *Information Theory and Statistics*. Peter Smith, 1978.
- [36] D. Labrousse, S. Dupont, and M. Berthod, "SAR interferometry: A Markovian approach to phase unwrapping with a discontinuity model," in *Proc. Int. Geoscience and Remote Sensing Symp.*, 1995, pp. 556–558.
- [37] R. Lagendijk, J. Biemond, and D. Boeke, "Identification and restoration of noisy blurred images using the expectation-maximization algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1180–1191, 1990.
- [38] J. Lee, K. Hoppel, S. Mango, and A. Miller, "Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 1017–1028, 1994.
- [39] J. Leitão and M. Figueiredo, "Interferometric image reconstruction as a nonlinear Bayesian estimation problem," in *Proc. 1995 IEEE Int. Conf. Image Processing*, 1995, vol. II, pp. 453–456.
- [40] J. Leitão and J. Moura, "Nonlinear phase estimation based on the Kullback distance," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, 1994, pp. 521–524.
- [41] ———, "Acquisition in phase demodulation: Application to ranging in radar/sonar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, pp. 581–599, 1995.
- [42] H. Lim, W. Xu, and X. Huang, "Two new practical methods for phase unwrapping," in *Proc. 1995 Int. Geosci. Remote Sensing Symp.*, 1995, pp. 196–198.
- [43] Q. Lin, J. Vesceky, and H. Zebker, "New approaches in interferometric SAR data processing," *IEEE Trans. Geosci. Remote Sensing*, vol. 30, pp. 560–567, 1992.
- [44] S. Madsen, "Spectral properties of homogeneous and nonhomogeneous radar images," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-23, pp. 583–588, 1987.
- [45] S. Madsen, H. Zebker, and J. Martin, "Topographic mapping using radar interferometry: Processing techniques," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 246–256, 1993.
- [46] J. Marroquin and M. Rivera, "Quadratic regularization functionals for phase unwrapping," *J. Opt. Soc. Amer. A*, vol. 12, pp. 2393–2400, 1995.
- [47] J. Marroquin, M. Tapia, R. Rodriguez-Vera, and M. Servin, "Parallel algorithms for phase unwrapping based on Markov random field models," *J. Opt. Soc. Amer. A*, vol. 12, pp. 2578–2585, 1995.
- [48] J. Moura, "Stochastic filtering: linear and nonlinear," in *Signal Processing*, J. L. Lacoume and R. Stora, Eds. Amsterdam, The Netherlands: North Holland, 1987.
- [49] J. Moura and A. Baggeroer, "Phase unwrapping of signals propagated under the Arctic ice crust: A statistical approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 617–630, 1988.
- [50] J. Moura and N. Balram, "Recursive structure of noncausal Gauss–Markov random fields," *IEEE Trans. Inform. Theory*, vol. 38, pp. 334–354, 1992.
- [51] S. Pandit, N. Jordache, and G. Joshi, "Data-dependent systems methodology for noise-insensitive phase unwrapping in laser interferometric surface characterization," *J. Opt. Soc. Amer.*, vol. 11, pp. 2584–2592, 1994.
- [52] M. Pritt, "Multigrid phase unwrapping for interferometric SAR," in *Proc. 1995 Int. Geoscience and Remote Sensing Symp.*, 1995, pp. 562–564.
- [53] C. Rathjen, "Statistical properties of phase-shift algorithms," *J. Opt. Soc. Amer. A*, vol. 12, pp. 1997–2008, 1995.
- [54] S. Song, S. Napel, N. Pelc, and G. Glover, "Phase unwrapping of MR phase images using Poisson equation," *IEEE Trans. Image Processing*, vol. 4, pp. 667–676, 1995.
- [55] U. Spagnolini, "2-D phase unwrapping and instantaneous frequency estimation," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 579–589, 1995.
- [56] A. J. Viterbi, *Principles of Coherent Communication*. New York: McGraw-Hill, 1971.
- [57] H. Vrooman and A. Maas, "Image processing algorithms for the analysis of phase-shifted speckle interference patterns," *Appl. Opt.*, vol. 30, pp. 1636–1641, 1991.
- [58] J. Woods, "Two-dimensional discrete Markovian fields," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 232–240, 1972.
- [59] J. Woods and C. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 473–482, 1977.
- [60] H. Zebker and R. Goldstein, "Topographic mapping from interferometric synthetic aperture radar," *J. Geophys. Res.*, vol. 91, pp. 4993–4999, 1986.



**José M. N. Leitão** (M'95) received the E.E. and Ph.D. degrees in electrical engineering in 1970 and 1983, respectively, and the "Agregado" degree in electrical and computer engineering in 1992, all from the Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal.

He was with the Laboratory of Physiology, Instituto Gulbenkian de Ciência, Oeiras, Portugal, from 1970 to 1972. After spending three years at the University of Tübingen, Germany, he joined the faculty of IST in 1976, where he is currently a Full

Professor with the Department of Electrical and Computer Engineering. He is also the coordinator of the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications. His main research interests are communication and information theory, pattern recognition, and signal and image processing.



**Mário A. T. Figueiredo** (S'87–M'95) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering, all from Instituto Superior Técnico (IST), Technical University of Lisbon, Portugal, in 1985, 1990, and 1994, respectively.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, IST. He is also a Researcher with the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications. His research interests include image processing, computer vision, and information and coding theory.