Phase Local Approximation (*PhaseLa*) Technique for Phase Unwrap From Noisy Data

Vladimir Katkovnik, Jaakko Astola, *Fellow, IEEE*, and Karen Egiazarian, *Senior Member, IEEE*

***Abstract*—The local polynomial approximation (LPA) is a nonparametric regression technique with pointwise estimation in a sliding window. We apply the LPA of the argument of cos and sin in order to estimate the absolute phase from noisy wrapped phase data. Using the intersection of confidence interval (ICI) algorithm, the window size is selected as adaptive pointwise varying. This adaptation gives the phase estimate with the accuracy close to optimal in the mean squared sense. For calculations, we use a Gauss–Newton recursive procedure initiated by the phase estimates obtained for the neighboring points. It enables tracking properties of the algorithm and its ability to go beyond the principal interval $[-\pi, \pi)$ and to reconstruct the absolute phase from wrapped phase observations even when the magnitude of the phase difference takes quite large values. The algorithm demonstrates a very good accuracy of the phase reconstruction which on many occasion overcomes the accuracy of the state-of-the-art algorithms developed for noisy phase unwrap. The theoretical analysis produced for the accuracy of the pointwise estimates is used for justification of the ICI adaptation algorithm.**

***Index Terms*—Adaptive window size, interferometric imaging, local polynomial approximation (LPA), phase image reconstruction, phase unwrapping.**

I. Introduction

A VARIETY OF imaging systems deal with phase measurements using coherent radiation in order to illuminate objects. The reflected scattered return carries information on the physical and geometrical properties of objects such as shape, deformation, structure of surface, and movement. Two-dimensional phase estimation has many important applications in different areas. For instance, in synthetic aperture radar interferometry, the phase value is proportional to a terrain elevation height; in magnetic resonance imaging, the phase is used to measure a magnetic field inhomogeneity. Other examples are in adaptive optics, diffraction tomography, nondestructive testing, deformation, and vibration measurements (e.g., [1]–[3]).

Common to these applications is that the observations are periodical functions of the phase which can be interpreted as the *principal* phase value, or *wrapped* phase, defined on the interval $[-\pi, \pi)$. Accordingly, it is impossible to unambiguously reconstruct the original, nonrestricted values, hereafter referred to as the absolute phase, unless additional assumptions are introduced. If an absolute phase value is outside the principal interval $[-\pi, \pi)$, the observed value is wrapped into this interval, corresponding to an addition or subtraction of an integer number of 2π. The wrapped ϕ and absolute phase φ are linked by the equation $\varphi = \phi + 2\pi k$, $\phi \in [-\pi, \pi)$, where k is integer. The wrapping operator is equivalent to division by module 2π, $\phi = \mathrm{mod}\{\varphi + \pi, 2\pi\} - \pi$, which separate φ on two parts: the fractional part ϕ and the integer part defined as $2\pi k$.

Many applications start from estimation of the phase for the principal interval and further extend these estimates to nonrestricted values. This last procedure is known as *phase unwrapping*. What makes this problem more difficult is that the measured values are usually corrupted by noise. The standard formulation of the noisy phase unwrapping starts from the observation model in the form

$$z_\phi = W(\varphi + \Delta\varphi) \tag{1}$$

where φ is the absolute phase, $\Delta\varphi$ is a random error additive to φ, and z_ϕ is the observed noisy wrapped phase. Here, W denotes a wrapping operator transforming the noisy absolute phase $z_\varphi = \varphi + \Delta\varphi$ to the interval $[-\pi, \pi)$.

Assume that the observations are given on the regular 2-D grid, $X = \{x, y : x = 1, 2, \ldots N_x, y = 1, 2, \ldots N_y\}$. The unwrapping problem is to reconstruct the absolute phase $\varphi(x, y)$ from the wrapped noisy $z_\phi(x, y)$ provided $x, y \in X$.

There is no one-to-one relation between the wrapped and unwrapped phase. Surprisingly, differentiation of the observations can resolve this ambiguity or at least to reduce it. Assume for a moment that there is no noise in observations, i.e., $z_\phi = \phi = W(\varphi)$.

Let Δ_x and Δ_y be difference operators on arguments x and y, respectively: $\Delta_x\varphi(x, y) = \varphi(x, y) - \varphi(x-1, y)$, $\Delta_y\varphi(x, y) = \varphi(x, y) - \varphi(x, y-1)$.

Proposition 1 [4]: Assume that the absolute phase φ satisfy to the conditions

$$-\pi \le \Delta_x\varphi(x, y) < \pi, \quad -\pi \le \Delta_y\varphi(x, y) < \pi \tag{2}$$

then

$$\begin{aligned} \Delta_x\varphi(x, y) &= W(\Delta_x\phi(x, y)) \\ \Delta_y\varphi(x, y) &= W(\Delta_y\phi(x, y)). \end{aligned} \tag{3}$$

According to Proposition 1, the phase $\varphi(x, y)$ can be restored by a two stage algorithm. First, the differences (derivatives) $\Delta_x\varphi(x, y)$ and $\Delta_y\varphi(x, y)$ are calculated according to the formula (3). Second, the phase $\varphi(x, y)$ is reconstructed by summation (integration) of these differences. It gives the phase estimate up to an additive constant. The result (3) applied to the

Manuscript received March 13, 2007; revised December 31, 2007. This work was supported by the Academy of Finland, Project No. 213462 (Finnish Centre of Excellence program 2006–2011). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Tamas Sziranyi.

The authors are with the Signal Processing Institute, University of Technology of Tampere, Tampere, Finland (e-mail: katkov@cs.tut.fi; jta@cs.tut.fi; karen.eguiazarian@tut.fi).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2008.916046

noise data (1) says that by differentiation of the noisy data and wrapping these derivatives, we obtain the derivatives of the unobserved noisy absolute phase $\varphi(x,y)$

$$\begin{aligned}\Delta_x(\varphi(x,y)+\Delta\varphi(x,y)) &= W(\Delta_x z_\phi(x,y))\\ \Delta_y(\varphi(x,y)+\Delta\varphi(x,y)) &= W(\Delta_y z_\phi(x,y)).\end{aligned} \quad (4)$$

Equation (2) can be treated as the Nyquist condition stating that the harmonic signal should be sampled at least twice for the period. In this case, there are no aliasing effects and the absolute phase can be reconstructed from the samples.

The techniques developed for phase unwrapping can be roughly separated in two large classes. The algorithms of the first class use the mentioned above two-stage approach with estimation of the gradient at the first stage and the following integration of this gradient at the second stage.

There are two main difficulties in this approach. First, the sampling conditions (2) often are not fulfilled for noisy data. Then the procedure cannot guarantee a correct phase reconstruction. Smoothness assumptions imposed on the absolute phase are used for regularization of the problem in order to improve the situation. Second, numerical differentiation as well as numerical integration are not trivial operation for noisy data. The differentiation results in increasing the noise level and the integration is an inverse ill-conditined operation also sensitive to noise. Thus, this two stage procedure should include filtering attenuating noise effects.

The algorithms of the second class are based on direct reconstruction of the absolute phase. Some of these algorithms are simple and unwrap the phase information by adding or subtracting 2π along the line and row whenever the phase difference between adjacent pixels is larger than 2π. However, abrupt phase changes in the absolute phase and experimental errors result in phase-unwrapping errors. Modified and more complex versions of the algorithm are based on modeling the absolute phase surface and include special tests on congruence of the phase estimate. A lot of methods are developed based on local and global phase modeling and test criteria.

A comprehensive review of the phase unwrap field is given in [1]. A recent advance in the area is reviewed in [5]–[9]. Further, we highlight briefly some of the basic methods and recent results in connection to the approach proposed in this paper.

A. Differentiation-Integration Methods

If the hypothesis (2) is not fulfilled, the integration of the gradient results are path dependent, i.e., the phase deviation between two points depends on the integration path linking these two points. Path following algorithms [10], [11] are developed for integrations over lines in the wrapped phase image where the Itoh condition (2) holds and the integration gives self-consistent results. In branch-cut methods [12], the integration paths are restricted by cuts, which cannot be crossed. These cuts are defined as the local inconsistencies calculated from the discrete derivatives.

An efficient solution to the unwrap problem is the minimum cost flow algorithm [13]. This algorithm is based on the consideration that when the condition (2) is violated the difference between the derivative of the absolute and wrapped phase is equal to multiples of 2π which should be added to the measured wrapped phase derivatives to achieve the absolute phase derivatives. The algorithm chooses these multiples by minimizing a global least-square norm criterion.

B. Direct Phase Fit

Another approach to 2-D phase unwrapping is based on mathematical formulation of the problem with reconstruction obtained by a constrained or nonconstrained global optimization. These methods are based on a least square estimate of the phase by minimizing the squared norm between the derivative estimate and unknown derivatives of the unwrapped phase [14], [15]. Nonquadratic norms L_p with $p<2$ also used for this sort of fitting [1], [16], [17] formulated as follows:

$$\begin{aligned}J = \sum_{x,y} |W(\Delta_x z_\phi(x,y)) - \Delta_x\varphi(x,y)|^p \\ + \sum_{x,y} |W(\Delta_y z_\phi(x,y)) - \Delta_y\varphi(x,y)|^p\end{aligned} \quad (5)$$

where $z_\phi(x,y)$ are given data and φ is the estimated absolute phase. Minimizing (5) yields a smooth phase reconstruction but may have large phase random error in the presence of noise and phase discontinuities [18]. This unwrapped phase usually fails the congruence test, which requires that rewrapping the unwrapped result reproduce the measured phase. In order to reduce these noise effects, a pixel-by-pixel weighting in (5) has been proposed [19].

The criterion

$$\begin{aligned}J = -\sum_{x,y}\lambda(x,y)\cos(\varphi(x,y)-z_\phi(x,y)) \\ +\frac{\mu}{2}\sum_{x,y}[v(x,y)|\Delta_x\varphi(x,y)|^2 + w(x,y)|\Delta_y\varphi(x,y)|^2]\end{aligned} \quad (6)$$

is minimized on φ in the $Z\pi M$ algorithm recently proposed in [5]. Here, the first summand is a fidelity term measuring a data-estimate divergence and the second summand is a penalty term imposing the smoothness conditions for the estimated absolute phase φ. The $Z\pi M$ algorithm is recursive with the unwrapping step minimizing the penalty term with respect to the integer k in $\varphi=\phi+2\pi k$ provided that the wrapped values of the phase are fixed. This step requires a discrete optimization implemented by the network programming technique. The smoothing step is minimization of (6) with respect to the wrapped values of the phase provided a given k. The level of the smoothing is controlled by the regularization parameter μ and the weights v and w, in particular, indicating the ares where the absolute phase may be discontinuous.

C. Energy Minimization

A general nonquadratic version of the penalty term from (6)

$$J = \sum_{x,y}[v(x,y)V(\Delta_x\varphi(x,y)) + w(x,y)V(\Delta_y\varphi(x,y))] \quad (7)$$

is proposed in [9] as a novel energy criterion for phase unwrapping. Here V is a nonquadratic loss function. Inserting in (7) $\varphi=\phi+2\pi k$ with ϕ given by observations the unwrapping is reduced to minimization of J on integer k. This complex combinatorial minimization is solved by the algorithm which is justified for both convex and nonconvex V. The criterion (7) is argu-

mented in [9] as a prior distribution for the first-order Markovian random field model for the absolute phase. For the quadratic V the criterion (7) is a typical choice appealing to the gaussian distribution. A motivation behind selection of the nonquadratic V is to make the solution minimizing J to be sensitive to discontinuities and irregularities in the absolute phase φ.

II. Proposed Approach

We start from calculation of $\cos/\sin$ functions of the observed wrapped phase values and replace the original wrapped phase observations ϕ by $\cos\phi$ and $\sin\phi$. Because $\cos\phi=\cos\varphi$ and $\sin\phi=\sin\varphi$, a difference between wrapped and unwrapped phases disappears and we use a fit of these transformed observations for the absolute phase reconstruction. The wrapped phase is discontinuous even for a continuous absolute phase. It is one of the reasons to work in the phase domain (using $\cos\phi$ and $\sin\phi$) instead of the original wrapped phase observations.

It is assumed in our approach that the absolute phase is a continuous function of the arguments x, y and allows a good polynomial approximation in a neighborhood of the estimation point. It is important that the size and possibly the shape of this neighborhood can be unknown and also is a subject of estimation.

In general, this approach is from the class of the nonparametric regression techniques. The algorithm developed in this paper is based on two independent ideas: *local approximation* for design of nonlinear filters (estimators) and *adaptation* of these filters to unknown smoothness of the spatially varying absolute phase. We use *local polynomial approximation* (LPA) for *approximation* and *intersection of confidence interval* (ICI) for *adaptation.*

In this paper, the LPA is applied for direct approximation of the absolute phase using a polynomial fit in a sliding window. The window size as well as the order of the polynomial define the accuracy of this approximation. The *window size* is considered as a *varying adaptation variable* of the algorithm.

The ICI is an *adaptation algorithm*. It searches for a largest local window size where the variance and the bias of the phase estimates are balanced. It is shown that the ICI adaptive LPA is efficient and allows to get a nearly optimal quality of estimation in particular for many image processing problems [20].

The polynomial modeling for the phase unwrap is a popular idea starting from the work [21], where it has been used for the global phase fitting. The efficiency of the local phase fitting is demonstrated in particular in [22], where the phase unwrapping appeared in connection with 2-D magnetic resonance imaging data. In the paper [23], the linear local polynomial approximation is developed for height profile reconstruction from multifrequency InSAR data. In the method called "local planes parameters estimation," the coefficients of the LPA are estimated by optimization of the likelihood criterion. Note that, in this paper, the efficient unwrap is achieved due to the multifrequency measurements.

Using the LPA fit for the phase unwrap based on the phase tracking is a main subject of papers [24] and [25].

The LPA and phase tracking developed in this paper are original mainly by the adaptive window size selection making the noise suppression more efficient and the risk of unwrapping error much lower.

There is a variety of phase observation models depending on measurement principals where the developed technique is applicable. Here, we wish to mention two basic ones.

1) $\cos/\sin$ *observations*

$$u_1=A\cos\varphi+n_1,\quad u_2=A\sin\varphi+n_2 \tag{8}$$

where A is the amplitude of the harmonic phase function, and n_1 and n_2 are noises. Then the wrapped phase ϕ is calculated according to the formulas

$$\cos\phi=\frac{u_1}{\sqrt{u_1^2+u_2^2}},\quad \sin\phi=\frac{u_2}{\sqrt{u_1^2+u_2^2}}. \tag{9}$$

2) *Phase-shifting observations*

$$u_l=A_0+A_1\cos(\varphi+\delta_l)+n_l,\quad l=1,\dots,L \tag{10}$$

where δ_l are fixed shifted phases, A_0 is a background intensity, A_1 is an amplitude of the harmonic phase function, and n_l are noises.

One of the most popular choices is $\delta_l=(l-1)\pi/2$, $l=1, 2, 3, 4$. Then, the intensities A_0, A_1 and the phase φ can be found with the phase defined according to the formulas

$$\cos\phi=\frac{u_1-u_3}{\sqrt{(u_4-u_2)^2+(u_1-u_3)^2}}$$
$$\sin\phi=\frac{u_4-u_2}{\sqrt{(u_4-u_2)^2+(u_1-u_3)^2}}. \tag{11}$$

A number of phase-shifting observations with different phase shifts and different number of observations are used in interferometric measurements [2, pp. 245–251].

All observation models similar to (9) and (11) can be represented in the form (1), where z_ϕ is a noisy wrapped phase and $\Delta\varphi$ is an error of the absolute phase φ. With random n_l in (8) and (10), and possibly random amplitudes A, the error $\Delta\varphi$ is random and in general phase dependent.

We assume that the observed data are already in the phase form (9). As the first step, we calculate

$$z_1=\cos(z_\phi),\quad z_2=\sin(z_\phi) \tag{12}$$

and call these variables *transformed noisy observations.* These noisy input data for the phase unwrap always can be represented as

$$z_1=\cos(\varphi+\Delta\varphi),\quad z_2=\sin(\varphi+\Delta\varphi) \tag{13}$$

where $\Delta\varphi$ denotes the error in the absolute phase φ caused by the observation errors in z_ϕ.

We apply LPA in order to approximate φ as an argument of the harmonic functions in (13). In principal, this idea can be exploited directly in the argument of the wrap operator W in (1). However, the wrap operator is discontinuous with respect to φ and use of the $\cos/\sin$ transform allows to replace it by the smooth differentiable one.

We call the proposed algorithm *PhaseLa* from "*phase local approximation*." The contribution of this paper is a development of this algorithm. Experiments show that this novel algorithm demonstrates a very good performance in comparison with some of the state-of-the-art techniques.

The rest of the paper is organized as follows. Section III introduces the idea and computational aspects of the LPA for the pointwise estimation and tracking the varying phase. The adaptive version of the LPA is introduced in Section IV, where the ICI rule is presented as the algorithm for the pointwise optimization of the window size. Overall, the *PhaseLa* algorithm organization is discussed in Section V. Simulation experiments analyzing the accuracy of the proposed algorithm are given in Section VI. The results are discussed in Section VII.

III. Phase LPA

Let us recall the basic ideas of LPA (e.g., [20]) and introduce LPA estimates of the phase. Assume that in some neighborhood of the point (x, y) the phase $\varphi(x, y)$ can be represented in the form

$$\tilde{\varphi}(x_s, y_s|\mathbf{c}) = \mathbf{p}^T(x_s, y_s)\mathbf{c} \tag{14}$$

where $\mathbf{p} = (p_1, p_2, p_3)^T$ is a vector of the first order polynomials $p_1 = 1, p_2 = x, p_2 = y$, and $\mathbf{c} = (c_1, c_2, c_3)^T$ is a vector of unknown parameters. The loss function of the local fit is defined as (15), shown at the bottom of the page.

The straightforward manipulations show that this expression is equivalent to

$$L_h(x, y, \mathbf{c}) = \sum_s w_{h,s}[1 - \cos(z_\phi(x + x_s, y + y_s) - \tilde{\varphi}(x_s, y_s|\mathbf{c}))] \tag{16}$$

and the fit parameter $\mathbf{c}$ is defined as a solution of the optimization problem

$$\hat{\mathbf{c}}(x, y) = \arg\min_{\mathbf{c}} L_h(x, y, \mathbf{c}). \tag{17}$$

The LPA estimates of the phase φ and the first derivatives $\varphi_x^{(1)}$, $\varphi_y^{(1)}$ are as follows [20]:

$$\begin{aligned} \hat{\varphi}(x, y) &= \hat{c}_1(x, y), \quad \hat{\varphi}_x(x, y) = \hat{c}_2(x, y) \\ \hat{\varphi}_y(x, y) &= \hat{c}_3(x, y). \end{aligned} \tag{18}$$

The window $w_{h,s}$ in (16) defines a set of neighborhood observations and their weights in estimation for x. The window size (*scale)* parameter h in w_h gives the size of the window and usually used in the form $w_h(x, y) = w(x/h, y/h)$, $h > 0$.

In particular, for the square uniform window $w = 1$ for $|x| \leq 1$, $|y| \leq 1$ and $w = 0$; otherwise, it means that $w_h = 1$ for $|x| \leq h$, $|y| \leq h$ and $w_h = 0$, otherwise. A smaller or larger h narrows or widens the window w_h, respectively.

The window function w_h can be symmetric or nonsymmetric with respect to the origin point $x = 0, y = 0$. It is assumed that the size of the support of w_h is larger then three (number of the parameters in $\mathbf{c}$ to be found).

The formula (18) shows that we obtain simultaneously the estimates of the phase $\hat{\varphi}$ and the instantaneous spatial frequencies $\hat{\varphi}_x$ and $\hat{\varphi}_y$. These estimates depend of the coordinate (x, y) and the window size h.

We wish to emphasize the nonparametric nature of the introduced estimates as the polynomial approximation (14) is used only for a single "central" point $x_s = y_s = 0$. For the phase, it gives $\hat{\varphi}(x, y) = \tilde{\varphi}(0, 0|\mathbf{c}) = \hat{c}_1(x, y)$, and for the derivatives $\hat{\varphi}_x(x, y) = \partial\tilde{\varphi}(x, y|\mathbf{c})/\partial x|_{x=0,y=0} = \hat{c}_2(x, y)$, $\hat{\varphi}_y(x, y) = \partial\tilde{\varphi}(x, y|\mathbf{c})/\partial y|_{x=0,y=0} = \hat{c}_3(x, y)$. The result of this pointwise use of LPA is that the parametric estimate (14) becomes nonparametric ones, i.e., $\hat{\varphi}(x, y)$ is a nonlinear with respect to x and y sometimes more depending on the data than on the order of the used approximation. All ideas of the standard LPA concerning the window w (shape, anisotropy, directionality, etc.), the scaling h (scalar, multivariate), and estimation of the signal and derivatives [20] are valid in the considered nonparametric pointwise estimation of the phase.

Here, we discuss the linear first order LPA as it is used in the forthcoming simulation experiments. A generalization to higher or lower degrees of polynomials in the model (14) or to the basis functions different from polynomials is straightforward.

A. Pointwise Estimate Calculation

Minimization of $L_h(x, y, \mathbf{c})$ nonquadratic with respect to $\mathbf{c}$ cannot be expressed in an analytical form and requires numerical recursive calculations using the vector-gradient: $\partial_{\mathbf{c}} L_h(x, y, \mathbf{c}) = (\partial_{c_i} L_h(x, y, \mathbf{c}))_{M\times 1}$, and the second derivative (*Hessian*) matrix: $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x, y, \mathbf{c}) = (\partial_{c_i}\partial_{c_j} L_h(x, y, \mathbf{c}))_{M\times M}$. Here we use M to denote the dimension of the vector $\mathbf{c}$ with $M = 3$ for the considered particular case.

A gradient descent recursive procedure for (17) has the standard form (e.g., [26]–[28])

$$\mathbf{c}^{(k+1)} = \mathbf{c}^{(k)} - \gamma_k \mathbf{A}^{(k)} \partial_{\mathbf{c}} L_h(x, y, \mathbf{c}^{(k)}), \quad k = 0, 1, \ldots \tag{19}$$

where $\mathbf{c}^{(k)}$ are successive iterations of $\mathbf{c}$, and the gradient $\partial_{\mathbf{c}} L_h$ is calculated for $\mathbf{c} = \mathbf{c}^{(k)}$.

The possible procedures are different by an $M \times M$ weight matrix $\mathbf{A}^{(k)}$ and a step size parameter $0 < \gamma_k \leq 1$.

1) *Simple gradient descent.* The identity matrix is used for $\mathbf{A}^{(k)}$, $\mathbf{A}^{(k)} = \mathbf{I}_{M\times M}$. The convergence rate is linear, $||\mathbf{c}^* - \mathbf{c}^{(k+1)}|| \leq q_k ||\mathbf{c}^* - \mathbf{c}^{(k)}||$, characterized by the parameter $q_k, 0 \leq q_k < 1$. Here, $\mathbf{c}^*$ is a vector of the optimal values of $\mathbf{c}$ minimizing $L_h(x, y, \mathbf{c})$. The step size parameter $\gamma_k = \gamma$ is selected in order to enable the convergence of the iterations, $q_k < 1$. A main drawback of the algorithm

$$\begin{aligned} &L_h(x, y, \mathbf{c}) \\ &\quad = \frac{1}{2}\sum_s w_{h,s}\{[z_1(x + x_s, y + y_s) - \cos\tilde{\varphi}(x_s, y_s|\mathbf{c})]^2 + [z_2(x + x_s, y + y_s) - \sin\tilde{\varphi}(x_s, y_s|\mathbf{c})]^2\} \\ &w_{h,s} = w_h(x_s, y_s) \geq 0 \end{aligned} \tag{15}$$

is a low convergence rate as q_k is close to 1 if the Hessian matrices $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h(x,\mathbf{c}^{(k)})$ are ill conditioned.

2) *Newton method.* The inverse Hessian matrix is used for $\mathbf{A}^{(k)}$

$$\mathbf{A}^{(k)} = (\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h(x,y,\mathbf{c}^{(k)}))^{-1}. \tag{20}$$

Here, we assume that $\mathbf{A}^{(k)}$ is inverse or pseudo-inverse of the $M \times M$ matrix $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h(x,\mathbf{c}^{(k)})$.The convergence is quadratic, $||\mathbf{c}^* - \mathbf{c}^{(k+1)}|| \le q_k||\mathbf{c}^* - \mathbf{c}^{(k)}||^2$, $q_k > 0$. This convergence rate is very good but the algorithm is sensitive with respect to initialization. For the quadratic convergence a good initial guess $\mathbf{c}^{(0)}$ is required.

3) *Gauss–Newton method.* The $\mathbf{A}^{(k)}$ in (19) is a special approximation of the inverse Hessian matrix. The convergence rate is linear but with small q_k. The convergence rate is comparatively insensitive with respect to the initialization.

In our experiments, we use the *Gauss–Newton* algorithm for calculation of the estimates as the most practically efficient one.

The straightforward manipulations give the vector-gradient and the Hessian matrix in the form

$$\begin{aligned}\partial_{\mathbf{c}}L_h &= \sum_s w_{h,s}\sin(z_\phi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c})) \\ &\quad \times \mathbf{p}(x_s,y_s) \qquad (21)\\ H &= \partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h \\ &= \sum_s w_{h,s}\cos(z_\phi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c})) \\ &\quad \times \mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s). \qquad (22)\end{aligned}$$

For the Gauss–Newton method, the matrix $\mathbf{A}$ in (19) is calculated as follows [28]. First, we produce linearization of $\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}+\delta\mathbf{c})$ and $\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}+\delta\mathbf{c})$ in (15) assuming that $\tilde{\varphi}(x_s,y_s|\mathbf{c}+\delta\mathbf{c}) = \mathbf{p}^T(x_s,y_s)\mathbf{c} + \mathbf{p}^T(x_s,y_s)\delta\mathbf{c}$, where $\delta\mathbf{c}$ is a small variation of $\mathbf{c}$

$$\begin{aligned}&\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}+\delta\mathbf{c}) = \\ &\quad \sin\tilde{\varphi}(x_s,y_s|\mathbf{c}) + \cos\tilde{\varphi}(x_s,y_s|\mathbf{c})\cdot\mathbf{p}^T(x_s,y_s)\delta\mathbf{c} \\ &\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}+\delta\mathbf{c}) = \\ &\quad \cos\tilde{\varphi}(x_s,y_s|\mathbf{c}) - \sin\tilde{\varphi}(x_s,y_s|\mathbf{c})\cdot p^T(x_s,y_s)\delta\mathbf{c}.\end{aligned}$$

Further, substitute these series in $L_h(x,y,\mathbf{c})$ given in the form (15), then the matrix corresponding to the Gauss–Newton method is calculated as

$$\hat{\mathbf{H}} = \partial_{\delta\mathbf{c}}\partial_{\delta\mathbf{c}^T}L_h = \sum_s w_{h,s}\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s) \tag{23}$$

and in (19) $\mathbf{A} = \hat{\mathbf{H}}^{-1}$. The formula (23) can be obtained from (22) assuming that the error approximation of $\cos(z_\phi(x+x_s,y+y_s))$ by $\tilde{\varphi}(x_s,y_s|\mathbf{c})$ is small.

The Hessian matrix (22) is useful to analyze the convexity of the criterion $L_h(x,y,\mathbf{c})$. For the noiseless case we have $z_\phi(x+x_s,y+y_s) = \varphi(x+x_s,y+y_s)$. Substituting these expressions in (22), we find that

$$\begin{aligned}\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h = \sum_s w_{h,s}\cos(\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c})) \\ \cdot\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s). \qquad (24)\end{aligned}$$

Let the polynomials $\mathbf{p}(x_s,y_s)$ be linear independent in the area where $w_{h,s} > 0$. It follows that the matrix $\sum_s w_{h,s}\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s)$ is positive definite. Then we may conclude for (24) that provided

$$|\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c})| < \frac{\pi}{2} \tag{25}$$

$\cos(\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c})) > 0$ and the matrix $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h$ is also positive definite. It proves that the criterion $L_h(x,y,\mathbf{c})$ is locally strongly convex and the convergence of the gradient style algorithm (19) can be guaranteed at least locally provided a proper selection of the matrix $\mathbf{A}$ and the step size parameter γ.

B. LPA Phase Unwrapping

The recursive algorithm (19) gives the estimate for any (x,y) provided that in the neighborhood of this point there is sufficient number of observations (x_s,y_s). With initialization independent for each point this is only a denoising algorithm which does not assume the phase unwrap. Let us use this pointwise estimator as an element of a more complex procedure with a special sequence of the estimation points (x,y) arranged with underlying intention to reconstruct, say, a continuous surface $\varphi(x,y)$.

A straightforward idea is to use for initialization the estimates already obtained for neighboring points. In particular, it can be a line-by-line sequence starting from the pixel (1,1) and going along the first line as $(1,2),(1,3),\ldots,(1,N_x)$, further the pixels of the second line $(2,1),(2,2),\ldots,(2,N_x)$, and in a similar way up to the last line $(N_y,1),N_y,2),\ldots,(N_y,N_x)$. In this way, we order all pixels of the phase image as the sequence $\{x^{(n)},y^{(n)}\}_{n=1,\ldots N_1N_2}$.

Let $\mathbf{c}^{(n)}(x^{(n)},y^{(n)}|\mathbf{c})$ be the estimate for the point $(x^{(n)},y^{(n)})$ provided that the recursive algorithm (19) is initiated by the vector $\mathbf{c}$. The proposed tracking phase unwrapping algorithm can be given in the following sequential form:

$$\mathbf{c}^{(n)} = \mathbf{c}^{(n)}(x^{(n)},y^{(n)}|\mathbf{c}^{(n-1)}) \tag{26}$$

$$\begin{aligned}\hat{\varphi}(x^{(n)},y^{(n)}) = c_1^{(n)}, \quad \hat{\varphi}_x(x^{(n)},y^{(n)}) = c_2^{(n)} \\ \hat{\varphi}_y(x^{(n)},y^{(n)}) = c_3^{(n)}, \quad n = 2,\ldots,N_x\cdot N_y. \qquad (27)\end{aligned}$$

Note that this recursive procedure includes the recursive pointwise estimator (19) as an imbedded one.

The procedure (26) is initiated by $\mathbf{c}^{(1)}$ for the first point $(x^{(1)},y^{(1)})$, i.e., we need to define the phase and the first two derivatives of the phase for this point. These values can be taken from original observations, from boundary conditions or as *a priori* information. It may be surprising, but this simple idea works and works very well combining two complementary important goals: noise suppression and reconstruction of continuous (or piece-wise continuous) absolute phase surface.

Our experiments show that the algorithm is successful provided that the absolute phase differences in the neighboring pixels are not large, mainly not larger than $0.5 \div 1$ radians. If the absolute phase differences are smaller the accuracy is very good, even for a high level of the random noise. Once more, note that the unwrapping property of the algorithm is appeared as a result of tracking the phase evolution from pixel to pixel.

IV. Spatially Adaptive LPA

A. Estimate Accuracy

Using a linearization of (13) for small $\Delta\varphi$, we can rewrite this model in the standard additive-error form

$$z_1 \simeq \cos\varphi + \varepsilon_1, \quad z_2 \simeq \sin\varphi + \varepsilon_2$$
$$\varepsilon_1 = -\Delta\varphi \cdot \sin\varphi, \quad \varepsilon_2 = \Delta\varphi \cdot \cos\varphi. \tag{28}$$

Let us derive the formula for the random phase-error $\Delta\varphi$. According to (9), we have for (13) $z_1 = u_1/\sqrt{u_1^2+u_2^2}$, $z_2 = u_2/\sqrt{u_1^2+u_2^2}$, where u_1 and u_2 are defined by (8). Using the Taylor series with respect to small n_1 and n_2, we find that

$$z_1 \simeq \cos\varphi + \frac{1}{A}\sin^2\varphi \cdot n_1 - \frac{1}{A}\sin\varphi\cos\varphi \cdot n_2$$
$$z_2 \simeq \sin\varphi - \frac{1}{A}\sin\varphi\cos\varphi \cdot n_1 + \frac{1}{A}\cos^2\varphi \cdot n_2.$$

Comparing these formulas with (28), we conclude that $\Delta\varphi$ in (13) is calculated as

$$\Delta\varphi = -\frac{1}{A}\sin\varphi \cdot n_1 + \frac{1}{A}\cos\varphi \cdot n_2. \tag{29}$$

With $E\{n_1\} = E\{n_2\} = 0$ and $var\{n_1\} = var\{n_2\} = \sigma^2$, we find for $\Delta\varphi$ that

$$E\{\Delta\varphi\} = 0, \quad \sigma_\varphi^2 = E\{(\Delta\varphi)^2\} = \frac{\sigma^2}{A}. \tag{30}$$

Thus, for a small level of the noise, we can assume that the random $\Delta\varphi$ in (13) and (28) is zero-mean with signal independent variance σ_φ^2 as defined in (30).

The estimation accuracy is characterized by the error between the absolute phase and the corresponding estimate: $e_\varphi(x,y) = \varphi(x,y) - \hat{\varphi}(x,y)$. This error is composed from the systematic (bias) and random components corresponding to the deterministic φ and the random noise ε, respectively.

The window size h is a crucial parameter for the accuracy of estimation. When the window size h is small, the LPA gives a good smooth fit of signals, but then fewer number of observations are used and the estimates are more variable and sensitive with respect to the noise. The best choice of h involves a trade-off between the bias and variance, which depends on the degree of the LPA, a sample period, the noise variance, and the derivatives of φ of the orders beyond the degree used in the LPA.

We present the accuracy analysis of the LPA estimates in order to illustrate these statements. We derive the formulas for the bias and the variance valid for small estimation errors. Further, we use these results in the algorithm for data-driven adaptive window size selection.

The bias of the estimate is a difference between the true signal and the expectation of the estimate: $E\{e_\varphi(x,y)\} = \varphi(x,y) - E\{\hat{\varphi}(x,y)\}$. Properties of φ should be specified in order to evaluate this error.

Let us assume that the phase is a continuous twice differentiable function. The finite Taylor series with the residual term in the Lagrange form gives

$$\begin{aligned}\varphi(x+\Delta x, y+\Delta y) &= \varphi(x,y) + \varphi_x(x,y)\cdot\Delta x + \varphi_y(x,y)\cdot\Delta y \\ &+ \varphi_{xx}(x+\lambda\Delta x, y+\lambda\Delta y)\cdot\frac{(\Delta x)^2}{2} \\ &+ \varphi_{yy}(x+\lambda\Delta x, y+\lambda\Delta y)\cdot\frac{(\Delta y)^2}{2} \\ &+ \varphi_{xy}(x+\lambda\Delta x, y+\lambda\Delta y)\Delta x\cdot\Delta y, \quad 0\le\lambda\le 1.\end{aligned} \tag{31}$$

We restrict our analysis to the class of smooth differentiable functions with bounded second derivatives

$$\max_{x_s,y_s\in U_h}(|\varphi_{xx}(x+x_s,y+y_s)|, \quad |\varphi_{yy}(x+x_s,y+y_s)| \quad |\varphi_{xy}(x+x_s,y+y_s)|) \le L_2(x,y) \tag{32}$$

where $L_2(x,y)$ is finite and U_h is a support of the window $w_{h,s}$ in (16).

Then it follows from (31) and (32) that, for any Δx and Δy

$$|\varphi(x+\Delta x,y+\Delta y) - \varphi(x,y) - \varphi_x(x,y)\cdot\Delta x - \varphi_y(x,y)\cdot\Delta y| \le L_2(x,y)\frac{(|\Delta x|+|\Delta y|)^2}{2}. \tag{33}$$

Proposition 2 (Pointwise Accuracy): Let the hypothesis (32) hold and the observation model be in the form (28)–(30). Assume also that the window function w is symmetric (even with respect to both arguments x and y) then the accuracy of the estimates (18) of the phase is defined as follows. For the bias

$$|E\{e_\varphi(x,y)\}| \le \frac{L_2(x,y)\sum_s w_{h,s}(|x_s|+|y_s|)^2}{2\sum_s w_{h,s}}. \tag{34}$$

For the variance

$$var\{e_\varphi(x,y)\} \simeq \frac{\sigma^2}{A^2}\frac{\sum_s w_{h,s}^2}{(\sum_s w_{h,s})^2}. \tag{35}$$

The proof of this proposition is given in Appendix.

Discussion of Proposition 2.

1) The bias of signal estimates is defined by the absolute values of the second derivatives (through L_2). It is interesting that the bias of the estimates does not depend on sin and cos functions. These formulas for the bias errors coincide with the ones derived for the linear LPA estimates in [20, Ch. 5]. Following the technique used in this book, the orders of the bias errors can be specified with respect to h. It can be shown that $\sum_s w_{h,s}(|x_s|+|y_s|)^2/(2\sum_s w_{h,s}) \sim 0(h^2)$. Smaller h results in a smaller bias error. It corresponds to the intuitively clear idea that, with a smaller window size, the LPA is able to give a better approximation to a smooth signal with a smaller error.

2) It can be shown that $\sum_s w_{h,s} \sim 0(h)$ and $\sum_s w_{h,s}^2 \sim 0(h)$. Then we obtain for the variance of the estimates that $var\{e_\varphi(x,y)\} \sim 0(1/h)$. Naturally, larger window means smaller variance.
3) This dependence of the bias and variance with respect to h says that the mean squares error $E\left\{e_\varphi^2(x,y)\right\} = |E\{e_\varphi(x,y)\}|^2 + var\{e_\varphi(x,y)\}$ has a minimum on h, which gives the optimal bias-to-variance balance with the best mean squared accuracy of estimation. The analysis of the optimization and varying optimal selection of h is one of the subjects discussed in detail in [20, Ch. 5].

B. Adaptive Window Size Selection

The theoretical analysis and experiments show that the efficiency of the local approximation estimates can essentially be improved provided a correct selection of the window size h. It can be varying or invariant but properly selected. In signal processing and statistics, window size selection is a subject of many publications exploiting different ideas and techniques.

Recently, a novel class of algorithms known under a generic name *Lepski's approach* has been introduced in statistics and shown to be efficient. These algorithms are proposed for the pointwise varying window size adaptive nonparametric estimation. One of the modification of this general approach named the intersection of confidence interval (ICI) algorithm is simple in implementation and found a number of efficient applications in image processing. Here, we explain the idea of this algorithm with reference for details to the book [20, Ch.6].

Let H be a set of the ordered window sizes $H = \{h_1 < h_2 < \cdots < h_J\}$. The estimates $\hat\varphi_h \equiv \hat\varphi(x^{(n)}, y^{(n)})$ (27) are calculated for all $h \in H$ and compared. The subscript h in the estimate emphasizes its dependence on h. A special statistic is exploited in order to identify the window size close to the optimal one. This statistic needs only the estimates $\hat\varphi_h$ and the variances of these estimates σ_h^2 both calculated for $h \in H$. Then the confidence intervals of these estimates are defined as

$$Q_h = \{\hat\varphi_h - \Gamma\cdot\sigma_h, \hat\varphi_h + \Gamma\cdot\sigma_h\} \tag{36}$$

where $\Gamma > 0$ is a parameter of the algorithm and σ_h is calculated according to (35), $\sigma_h^2 \equiv var\{e_\varphi(x,y)\}$.

The ICI rule defines the adaptive window size denoted h^+ as the largest h of those in H, with the estimate which does not differ significantly from the estimates corresponding to the smaller window sizes. In order to identify this adaptive h^+, the successive intersection of the confidence intervals Q_h is considered starting from Q_{h_1} and Q_{h_2}. Specifically, *the pairwise intersection of the intervals Q_{h_j}, $1 \le h_j \le h_i$, is considered with increasing h_i . Let h^+ be the largest of those h_i for which the intervals Q_{h_j}, $1 \le h_j \le h_i$, have a point in common. This h^+ defines the desired adaptive window size and the adaptive estimate as $\hat\varphi_{h^+}$.*

For the varying pointwise adaptive estimation, these calculations are produced for all points (pixels). In the implementation, the ICI algorithm is used when the estimates for all points (x,y) are already calculated for all h. Then the algorithm works as a selector of the proper window size estimate for each point.

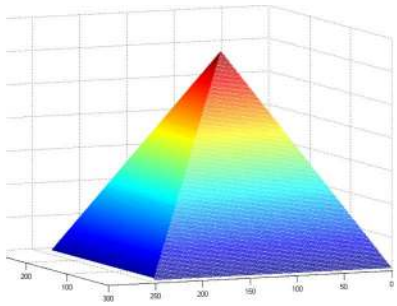


Fig. 1. Pyramid test function.

It is emphasized that the ICI adaptive window size enables values close to the optimal ones minimizing the mean squared error. However, Γ is an important parameter of the algorithm controlling the bias-variance balance in the estimate. Smaller Γ means a shift of this balance in favor of the bias, as smaller Γ results in smaller bias of the estimate. Contrary to it, larger Γ means a shift in favor of the variance, as larger Γ results in smaller variance of the estimate but possibly larger bias.

V. *PhaseLa* Algorithm

The variables z_1, z_2 defined by (12) are input signals of the *PhaseLa* algorithm.

Initialization of the vector $\mathbf{c}$

$$\begin{aligned}\mathbf{c} &= \mathbf{c}^{(1)} : c_1^{(1)} = \phi(x^{(1)}, y^{(1)})\\ c_2^{(1)} &= W(\Delta_x\phi(x^{(1)}, y^{(1)})), \quad c_3^{(1)} = W(\Delta_y\phi(x^{(1)}, y^{(1)}))\end{aligned} \tag{37}$$

where ϕ is the observed wrapped phase.

For every pixel of the sequence $(x^{(n)}, y^{(n)})$, $n = 2, \ldots, N_xN_y$:

1) **calculate the vectors** $c^{(n)}$ **and the point-wise estimates** $\hat\varphi_h^{(n)}$ **according to** (19), (23), and (26)–(27);

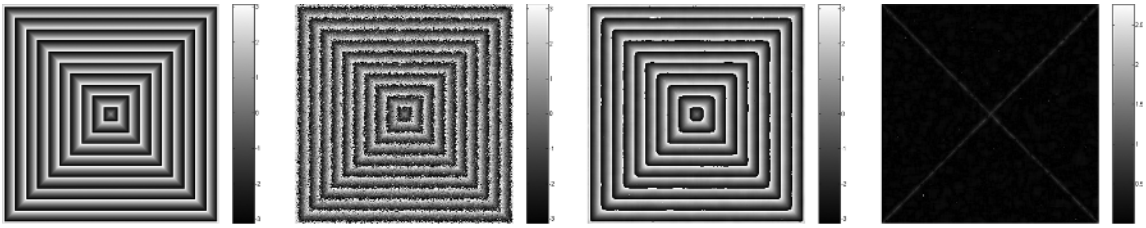
2) **repeat these calculations for all** $h = h_1, h_2, \ldots$;

3) **apply the ICI rule for selection of the best window size and the adaptive window size estimate** $\hat\varphi_{h^+}^{(n)}$.

The initialization (37) by the observed wrapped phase values is used only for the first pixel $(x^{(1)}, y^{(1)})$. For further pixels, the initiation is produced using the adaptive window size estimates obtained for the neighboring pixels where these estimates are already calculated according to the recursive procedure (27). The estimates for different h are calculated with the same initialization common for every particular pixel.

VI. Simulation Experiments

We are focused on simulated data in order to be able to evaluate the algorithm performance accurately. The observation models used in the experiments are described in detail. As the accuracy criterion we use the root-mean-squared-error,



(a) (b) (c) (d)

Fig. 2. Wrapped Pyramid phase: (a) true absolute, (b) noisy, (c) *PhaseLa* rewrapped, and (d) absolute errors between the true absolute and *PhaseLa* unwrapped phases.

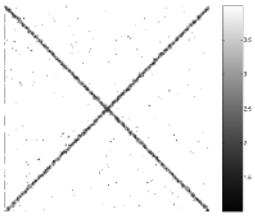


Fig. 3. ICI adaptive window sizes for Pyramid phase.

RMSE $= \sqrt{(1/N_xN_y)\sum(\varphi(x_s,y_s)-\hat{\varphi}(x_s,y_s))^2}$. The LPA is exploited with the uniform square windows w_h defined on the integer grid $U_h = \{x,y : x = -h,-h+1,\ldots,0,\ldots,h-1,h,y = -h,-h+1,\ldots,0,\ldots,h-1,h\}$.

Mainly, the ICI algorithm parameter $\Gamma = 2.0$ and the set of the window sizes $H = [1,2,3,4]$.

As a benchmark for comparison, we use the results obtained by the $Z\pi M$ algorithm [5], which is considered as one of the best algorithms developed for noisy data. We produce our experiments using the first order polynomial model. The derivative estimates play important role in unwrapping as the initialization of the recursive pointwise estimates includes also the initialization for derivative estimates. In this way, the phase tracking essentially exploits the continuity of the phase.

For all experiments, we use the Matlab codes of the *PhaseLa* algorithm available at http://www.cs.tut.fi.

A. Pyramidal Phase

The pyramidal absolute phase test function (Fig. 1) is defined by the formulas

$$\varphi = 0.5\cdot\min(\varphi_1,\varphi_2,\varphi_3,\varphi_4),\quad \varphi_1 = x$$
$$\varphi_2 = y,\quad \varphi_3 = 255-x,\quad \varphi_4 = 255-y$$

on the integer grid $x = (0:255)$, $y = (0:255)$. The maximum of φ is equal to 63.5 radians and the maximum of the pixel-wise difference is .5 radians.

In Fig. 2, one can see the wrapped true absolute phase φ, the noisy wrapped phase calculated according to the formulas (8)–(9), and the rewrapped phase reconstruction $W(\hat{\varphi})$. Comparing the wrapped true absolute phase and $W(\hat{\varphi})$, one may conclude that the filtering and unwrapping are quite accurate.

The adaptive window sizes shown in Fig. 3 give insight how the adaptation works. Mainly, the largest window size is selected excluding the areas near pyramid edges, where the adaptive window size takes the minimum value. In this way, the algorithm enables the maximum smoothing of the noise for the flat surfaces where the used linear model perfectly fits to the surface and the maximum window size can be used. For the edges, small window size allows to avoid the surface oversmoothing, however, at the price of a higher level of random errors. The effects of the varying window selection is illustrated also by the last image in Fig. 2 showing the absolute errors of the phase reconstruction. These errors are minimal on flat surfaces of the pyramid where the window sizes take maximal values and these errors are maximal along the pyramid edges where the adaptive window sizes are minimal.

TABLE I
RMSE FOR THE *PHASELA* AND $Z\pi M$ ALGORITHMS, PYRAMID TEST FUNCTION

Algorithm \ σ	.1	.2	.3	.4	.5
PhaseLa, $h=1$	.039	.071	.109	.152	.200
PhaseLa, $h=2$	.048	.059	.074	.098	.122
PhaseLa, $h=3$	.073	.077	.082	.095	.108
PhaseLa, $h=4$	.214	.201	.185	.112	.119
PhaseLa, *ICI*	**.029**	**.054**	**.075**	**.095**	**.113**
$Z\pi M$	.058	.084	.117	.154	.190

Numerical evaluation of the algorithm performance is illustrated in Table I. It shows the results for *PhaseLa* with invariant values of h =1, 2, 3, 4 and with ICI varying adaptive ones. The results are given for different values of the additive zero mean gaussian noise in (8). We can see in this table a difference between the estimates with invariant h and varying adaptive one in the row corresponding the ICI adaptive algorithm. In all cases, the adaptive algorithm enables minimization of RMSE values and even slightly better results that the best one achieved for the invariant window size.

We also show the results given by the $Z\pi M$ algorithm (ten iterations). Comparing these results versus *PhaseLa* with the adaptive window size selection we may conclude that this adaptive algorithm gives a valuable improvement of the accuracy. RMSE values are about 1.5 times better for the *PhaseLa* algorithm than those for the $Z\pi M$ algorithm.

B. Ramp Phase

For the linear (*ramp*) absolute phase, the *PhaseLa* algorithm demonstrates the perfect performance as the ICI adaptation automatically selects the largest window size and in this way enables the best noise attenuation giving the unbiased estimate

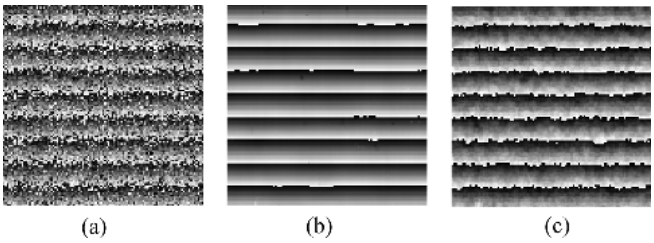


Fig. 4. Wrapped ramp phase: (a) noisy, (b) *PhaseLa* rewrapped, and (c) $Z\pi M$ rewrapped.

TABLE II
RMSE FOR THE *PHASELA* AND $Z\pi M$ ALGORITHMS, RAMP TEST FUNCTION

Algorithm \ σ	.1	.2	.3	.4	.5	.7	1.0
PhaseLa, $h=3$	.015	.03	.044	.060	.074	.115	.175
PhaseLa, $h=5$	.010	.019	.028	.039	.048	.072	.110
PhaseLa, $h=7$	.008	.013	.021	.028	.034	.053	.077
PhaseLa, $h=9$	.006	.012	.018	.025	.032	.046	.065
PhaseLa, ICI	**.006**	**.012**	**.018**	**.025**	**.032**	**.047**	**.066**
$Z\pi M$	.050	.078	.115	.147	.186	.255	.325

of the phase. In these experiments, we use larger values of the window sizes, $H=[3,5,7,9]$ and $\Gamma=5.0$.

The ramp function is defined as $\varphi(x,y)=0.5\cdot x$ for $x=(0:127)$, $y=(0:127)$. Thus, the maximum value of φ is 63.5 with the maximum difference between pixels equal to .5. The numerical results are shown in Table II for different noise standard deviation σ in the observation model (8). Comparing *PhaseLa* versus the $Z\pi M$ algorithm is definitely in favor of *PhaseLa*. Fig. 4 illustrates what this difference in RMSE values means visually. These images are given for $\sigma=1$. The adaptive *PhaseLa* nearly perfectly suppresses the noise in this heavy noisy data with only a few erroneous pixels clear seen in the image.

C. Parabolic Phase

In these experiments, we use the model studied in [18][1] The phase φ is a parabola defined by the formula

$$\varphi=2\pi\cdot\bar{\varphi},\quad \bar{\varphi}=3-\frac{(x-64)^2}{80}-\frac{(y-64)^2}{80}$$

where $\bar{\varphi}$ is a normalized phase to be estimated.

The level of the additive gaussian noise is characterized by the signal-to-noise ratio SNR $=20\log_{10}(A/\sigma)$. The observation model has a form (8) with $A=1$. The maximum values of φ and the phase difference are 18.85 and 2.28, respectively.

It is shown in [18] that the algorithm developed by the authors of this paper and the algorithm by Chen and Zebker [29] demonstrate nearly identical results which are much better than those obtained by the least square method sensitive with respect to noise.

[1]The model and conditions of this experiment important for comparison with the results in [18] are due to personal communication with L. Ying.

TABLE III
RMSE VALUES OBTAINED BY THE ALGORITHMS: PROPOSED IN [18], $Z\pi M$ AND *PHASELA*

SNR	Ying et al. [18]	$Z\pi M$	*PhaseLa*
8	.166	.027	**.019**
10	.128	.023	**.015**
12	.102	.020	**.013**
14	.077	.017	**.012**
16	.067	.016	**.010**
18	.050	.015	**.008**
20	.045	.014	**.008**

TABLE IV
RMSE VALUES OBTAINED BY THE ALGORITHMS $Z\pi M$ AND *PHASELA* FOR COHERENCE α

Algorithm \ α	.7	.75	.8	.85	.9	.95	.99
PhaseLa, $h=2$	.26	.23	.20	.18	.16	.13	.11
PhaseLa, $h=3$	3.31	10.27	.23	.22	.21	.20	.20
PhaseLa, $h=4$	7.73	6.78	7.7	8.2	5.9	4.3	2.53
PhaseLa, $h=5$	8.17	8.04	9.4	6.1	3.4	5.1	2.60
PhaseLa, ICI	**.25**	**.23**	**.21**	**.19**	**.17**	**.15**	**.11**
$Z\pi M$	2.54	1.18	.27	.24	.21	.17	.11

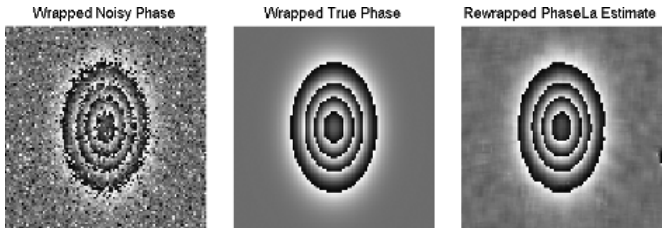


Fig. 5. Noisy data, wrapped true phase, and rewrapped phase reconstruction obtained by the *PhaseLa* algorithm ($\alpha=0.8$).

In Table III, we show RMSE values for: the best results from [18], and the results obtained by the *PhaseLa* and $Z\pi M$ algorithms. In this competition the $Z\pi M$ algorithm shows much better accuracy than that in [18] and the *PhaseLa* algorithm enables even better accuracy. Comparing with the $Z\pi M$ algorithm we can see that the accuracy of the *PhaseLa* algorithm is about 1.5 times better for all SNR.

D. InSAR Model

Here, we use the interferometric synthetic aperture radar (InSAR) model as it is introduced in [5] and follow assumptions and parameters discussed in details in this paper. The observed InSAR data are given by complex variables

$$x_1=a_1\exp(-j\varphi_1)+n_1,\quad x_2=a_2\exp(-j\varphi_2)+n_2 \quad (38)$$

where a_1, a_2 are complex valued amplitudes of the harmonic phase signals and n_1, n_2 are complex-valued observation errors.

All these variables are random independent zero-mean gaussian. The phase shift $\varphi=\varphi_1-\varphi_2$ is a parameter of

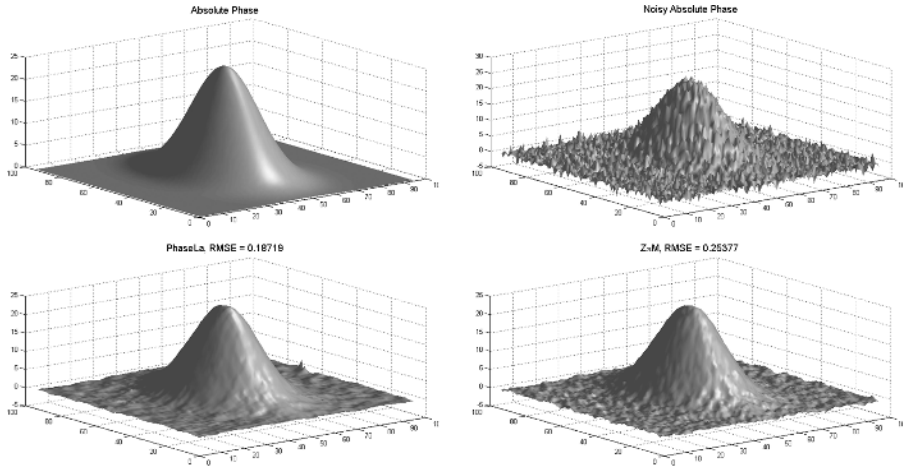


Fig. 6. True absolute phase, hypothetical noisy absolute phase, *PhaseLa* and $Z\pi M$ reconstructions ($\alpha = 0.8$).

interest. It is assumed that $\theta^2 = E\{|a_1|^2\} = E\{|a_2|^2\}$ and $\alpha = E\{a_1 a_2^*\}/\theta^2$ is real.

The input for the signal processing is calculated as the product $q = x_1 x_2^*$. If there is no noise, $n_1 = n_2 = 0$, we have

$$q = a_1 a_2^* \exp(-j\varphi) = |a_1||a_2| \exp(-j(\varphi + \Delta\varphi))$$

where the random error in the phase $\Delta\varphi$ is a phase difference of the random phases of a_1 and a_2.

Being rewritten in the form (8), it gives $u_1 = |a_1||a_2| \cos(\varphi + \Delta\varphi)$, $u_2 = |a_1||a_2| \sin(\varphi + \Delta\varphi)$ and further for the wrapped phase ϕ

$$z_1 = \cos\phi = \frac{u_1}{\sqrt{u_1^2 + u_2^2}} = \cos(\varphi + \Delta\varphi)$$
$$z_2 = \sin\phi = \frac{u_2}{\sqrt{u_1^2 + u_2^2}} = \sin(\varphi + \Delta\varphi). \qquad (39)$$

In this way, we arrive to the model (13) used as a starting point of our algorithm. It follows from Proposition 2 that the variance of the estimates depending on $A = |a_1||a_2|$ becomes random. If the noises n_1 and n_2 are nonzero the situation becomes even more complex with the amplitude A random and depending on the unknown absolute phase φ.

The $Z\pi M$ algorithm is compared in [5] versus a number of prominent algorithms proposed for phase unwrapping. Those of these algorithms which are developed for noiseless data are considered with a special prefiltering of the observed noisy wrapped phase. It is shown (see Table I in [5]) that the $Z\pi M$ algorithm with simultaneous smoothing and unwrapping demonstrates a great deal of advantage over all compared algorithms when the phase unwrapping is produced from noisy data. This fact gives a reason to compare in this paper the *PhaseLa* algorithm versus the $Z\pi M$ algorithm only as it is the best algorithm at least in the group studied in [5].

According to [5], the following is assumed. The absolute phase is gaussian $\varphi = A_\varphi \exp\left(-x^2/\left(2\sigma_x^2\right) - y^2/\left(2\sigma_y^2\right)\right)$, $\sigma_x = 10$, $\sigma_y = 15$, $A_\varphi = 14\pi$, with integer arguments x, y, $-49 \le x, y \le 50$. The amplitudes a_1 and a_2 are random with the variance $\theta = 1$ and $n_1 = n_2 = 0$ in (38). The coherence α is a varying parameter of simulation experiments.

The maximum values of the absolute phase is equal to 14π with the maximum value of the differences about 2.5 radians. For the considered noisy data, the phase difference is often takes values close to 2π.

Smaller and larger values of α correspond respectively to larger and smaller noise level in observations starting from $\alpha = 0.7$ means a very high intensity of the noise and going up to $\alpha = .99$ corresponding to nearly ideal noiseless data. The RMSE values are shown in Table IV for different values of the coherence α.

The *PhaseLa* algorithm compared versus the $Z\pi M$ algorithm mainly demonstrates a better performance. This advantage is very impressive for the high noise level with $\alpha = 0.7, 0.75$, where the $Z\pi M$ algorithm fails what follows from very large values of RMSE while the *PhaseLa* algorithm gives a reasonable accuracy of phase unwrapping. For a lower level of the noise ($\alpha \to 1$), the accuracy of the compared algorithms becomes close with the negligible difference for $\alpha = 0.99$. The quality of the *PhaseLa* method for noisy data ($\alpha = 0.8$) is illustrated in Fig. 5, where one can observe the noisy observations and the rewrapped phase estimate $W(\hat{\varphi})$. It is seen that visually the estimate is quite good. The 3-D imaging in Fig. 6 gives further illustrations. One can see here the true phase and what we call "*hypothetical noisy true phase*." The last signal is obtained by unwrapping the noisy wrapped observations z_ϕ and used only to give an idea what kind of a noisy signal, $\varphi + \Delta\varphi$, corresponds to z_ϕ. We also show the reconstructions obtained by the *PhaseLa* and $Z\pi M$ algorithms. The *PhaseLa* estimate is smoother and as followed from the smaller value of RMSE is more accurate than that for the $Z\pi M$ algorithm.

A distribution of the ICI adaptive window for $\alpha = 0.8$ is illustrated in Fig. 7.

It is important to emphasize a crucial role of adaptive window size selection in this experiments. Table IV shows that the unwrap using the LPA with a fixed window size fails with h =4, 5 for all α, and with $h = 3$ for $\alpha = 0.7, 0.75$. Nevertheless, we can see that the unwrap with the ICI adaptive window size is successful with a good accuracy. It says, that the ICI treats the fails in unwrapping as large bias errors. In this way, the ICI rule filters out the estimate with errors in unwrapping. However, it

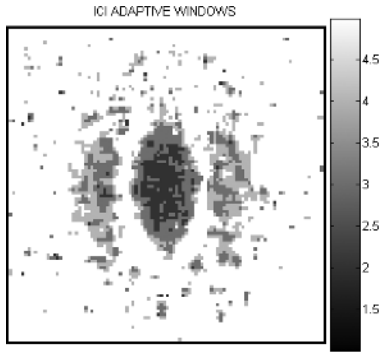


Fig. 7. Adaptive window sizes for the *PhaseLa* phase unwrap ($\alpha = 0.8$).

works correctly, if the estimates with different h starts from the properly unwrapped estimates.

VII. Concluding Remarks

This paper presents an efficient approach for absolute phase reconstruction from noisy wrapped phase measurements. The local approximation technique is exploited for the phase estimation and attenuation of noise effects. The unwrapping is achieved by successive phase reconstruction for neighboring pixels. The window size adaptation enables a reasonable compromise between the noise smoothing and preservation of details in phase image.

In what follows in this section, we discuss some principal and technical issues concerning our approach.

1) Overall, the developed LPA technique is the nonlinear least square method with a pointwise estimation in a sliding window. It can be treated also as a nonlinear recursive filter tracking (from pixel-to-pixel) phase values. To the best of our knowledge, this novel recursive filter is essentially different from recursive and nonrecursive procedures which have been used before now for noisy phase unwrap.
   In [1], the filtering is considered as a preprocessing preceding the main unwrapping algorithm. It is recommended in [1, Ch. 3] to filter independently two signals $\cos\phi$ and $\sin\phi$ with following recalculation of the wrapped phase values through these filtered $\cos\phi$ and $\sin\phi$. Our filtering is different because the local approximation is used directly for the reconstructed absolute phase as the argument of $\cos / \sin$ functions. In this way, the observation model is thoroughly exploited in the developed estimator with naturally a much more efficient filtering.
   If we compare our algorithm versus the Kalman-Busy style filter proposed in [30], we may note, first, that the filter in [30] is applied to observations given in the form (8), where the noise is additive. Our algorithm starts from the wrapped phase data and then there is no additive noise in the observations (13). Recall, that the observation noise is essential for the Kalman–Busy technique where no noise is a singular situation. Thus, different observation models and, as a result, a different setting of the problem are considered. The accuracy control imbedded in the recursive pointwise estimation in (19) and the window size adaptation makes a difference between our algorithm and the algorithm from [30] even deeper.
2) In this paper, we treat the $Z\pi M$ algorithm as a bench mark and use it for comparison. We show that on many occasion our algorithm demonstrates a better accuracy. It is interesting to discuss a difference between the algorithms.
   2.1) The $Z\pi M$ algorithm is a procedure with a solution obtained by minimization of the global (defined over whole image) criterion (6). The smoothness of the reconstructed phase φ is defined by the parameter μ. With $\mu \to 0$ there is no smoothness constraints at all and any $\varphi(x,y) = z_\phi(x,y) + 2\pi k$ gives the minimum value $J \to_{\mu \to 0} 0$. For large $\mu \to \infty$, the corresponding solution approaches a constant value as the phase differences should go to zero in order J will be bounded. Note that, in the *PhaseLa* algorithm, the estimate with an increasing window size gives the estimate which is linear with respect to the argument (x,y). Thus, the zero-order polynomial approximation is used in the $Z\pi M$ algorithm and the first order in the *PhaseLa* algorithm.
   Using the parameter μ and weight $\lambda(x,y)$ in (6), we can vary the smoothness of the solution and generate a variety of versions of the unwrapped absolute phase. In our simulation experiments, we assume that the parameters of the $Z\pi M$ algorithm are fixed as they are given in the author's code and show that in this case the *PhaseLa* algorithm demonstrates more accurate results. It is quite possible that there exists such tuning of μ and $\lambda(x,y)$ that the $Z\pi M$ algorithm performs better than the *PhaseLa* algorithm.
   However, variations of the weight $\lambda(x,y)$ can result in global changes of the phase φ, and it is a nontrivial task to enable the desirable pointwise smoothness correction through the solution of the global optimization. Contrary to $Z\pi M$, the *PhaseLa* procedure is local minimizing the local criterion (17) in the pointwise manner. In this way, the size of the estimation window is an efficient instrument for precise and straightforward control of the smoothness for every pixel individually. The ICI algorithm gives a rule to select a reasonable distribution of the window sizes over the phase image. Thus, locality and globality is the first issue differs the discussed two algorithms.
   2.2) The unwrapping is a key point of the $Z\pi M$ algorithm while the *PhaseLa* algorithm is focused on approximation and noise suppression. The minimization of J over integer k in (6) produces a global unwrapping.
   In the *PhaseLa* algorithm, the unwrapping is a result of the accurate approximation and careful fusing of the estimates for neighboring pixels. It means that for unwrapping we use the local analysis of the estimates only. The experiments confirms that this idea works well.
   2.3) Concerning the complexity of the $Z\pi M$ and *PhaseLa* algorithm we wish to note that the computation time of these algorithm is more less the same

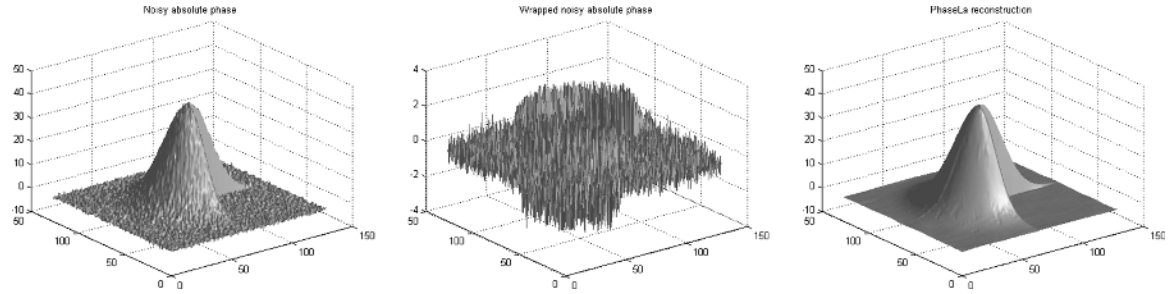


Fig. 8. Discontinuous absolute phase: noisy data and *PhaseLa* reconstruction.

provided that four iterations are used in the $Z\pi M$ algorithm. Larger number of iterations naturally mean longer computation time.

3) Using the adaptive window size for the phase unwrap is originated in our conference paper [31]. In this paper, the tracking of the phase is produced with a fixed window size. Thus, we obtain the estimates calculated with invariant window sizes and the ICI is used in order to select the best estimate for each pixel.
   In the *PhaseLa* algorithm presented in this paper, the phase tracking is performed on the estimates with already adaptive window sizes. Comparison of the algorithms is definitely in favor the *PhaseLa* algorithm which demonstrates much better performance.
4) The line-by-line phase restoration implemented in the *PhaseLa* algorithm is a not universally best strategy. In particular, the tracking mimicking path-dependent integration methods with local phase congruence tests can give a further improvement of the algorithm.
5) The LPA and the ICI procedures as they are presented in this paper are proposed for continuous and differentiable phase functions.

However, the LPA with the adaptive window size selection allows a number of modifications for more complex problems with nondifferentiable and discontinuous functions.

Let $\varphi(x,y)$ be a piece-wise continuous differentiable phase function. It means that the area I where this function is defined can be segmented on nonoverlapping subareas I_r, $\cup_{r=1}^{L} I_r = I$, $I_r \cap I_{r'} = \phi$ if $r \neq r'$, such that for any (x,y) exist I_r where $\varphi(x,y)$ is continuous and differentiable. Introduce the indicator (mask) function for I_r subarea, $M_r(x,y) = 1$ for $(x,y) \in I_r$ and $M_r(x,y) = 0$ otherwise. Assume that this segmentation is given. The *PhaseLa* is applicable for the phase unwrap in I_r provide that the weight $w_{h,s}$ in (16) are replaced by $w_{h,s} \cdot M_r$ and the algorithm is initiated by the data from this area. Fig. 8 illustrates the work of the algorithm in this situation. There are two subareas where the considered absolute phase is continuous. In one of these areas, it the gaussian density while in the second subarea (quadrant sector) the phase function is equal to zero. Fig. 8 shows the noisy absolute phase, the observed wrapped phase and the *PhaseLa* reconstructed unwrap phase. The algorithm demonstrates a very good performance.

To deal with nonsmooth functions when the piece-wise segmentation is unknown the adaptive anisotropic LPA can be applied. In this concept, the symmetric square window function $w_{h,s}$ is replaced by four/eight sectorial windows with the ICI window size selection independent for each sector. The final estimate is obtained by aggregation of the sectorial ones. These anisotropic estimates are highly sensitive with respect to discontinuity and anisotropic behavior of the reconstructed functions. This sort of methods in applications for image processing are discussed in [20, Ch. 7–8].

APPENDIX

Proof of Proposition 2: The minimum condition for unconstrained optimization (17) has a form $\partial_{\mathbf{c}} L_h(x,y,\hat{\mathbf{c}}) = 0$, where $\hat{\mathbf{c}}$ is a vector-estimate. Using the first two term of the Taylor series this equation gives

$$\partial_{\mathbf{c}} L_h(x,y,\hat{\mathbf{c}}) = \partial_{\mathbf{c}} L_h(x,y,\mathbf{c}^*) + \partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x,y,\mathbf{c}^*)\Delta\mathbf{c} = 0 \tag{40}$$

where $\mathbf{c}^* = (\varphi(x,y), \varphi_x(x,y), \varphi_y(x,y))^T$ is a vector of true values of the phase and the derivatives, $\Delta\mathbf{c} = \mathbf{c}^* - \hat{\mathbf{c}}$.

The vector gradient $\partial_{\mathbf{c}} L_h$ and the Hessian matrix $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h$ are defined in (21) and (22). Let us calculate the expectation of the Hessian matrix. Using (28), we have $E\{z_1\} \simeq \cos\varphi$ and $E\{z_2\} \simeq \sin\varphi$ and then

$$\begin{aligned} &E\{\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x,y,\mathbf{c}^*)\} \\ &\quad\simeq \sum_s w_{h,s}[\cos\varphi(x+x_s,y+y_s)\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) \\ &\qquad\qquad + \sin\varphi(x+x_s,y+y_s)\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}^*)] \\ &\qquad\times \mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s) \\ &\quad= \sum_s w_{h,s}\cos(\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c}^*)) \\ &\qquad\times \mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s). \end{aligned} \tag{41}$$

According to (33)

$$|\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c}^*)| \leq L_2(x,y)\frac{(|x_s|+|y_s|)^2}{2}. \tag{42}$$

For a small h we have $|\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c}^*)| \simeq 0$ and $\cos(\varphi(x+x_s,y+y_s) - \tilde{\varphi}(x_s,y_s|\mathbf{c}^*)) \simeq 1$ then

$$E\{\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x,y,\mathbf{c}^*)\} \simeq \sum_s w_{h,s}\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s) > 0. \tag{43}$$

For an increasing number of samples in U_h, there is a convergence in probability

$$\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x,y,\mathbf{c}^*) \to_P E\{\partial_{\mathbf{c}}\partial_{\mathbf{c}^T} L_h(x,y,\mathbf{c}^*)\} \simeq \sum_s w_{h,s}\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s).$$

$$E\{\Delta\mathbf{c}\} \simeq \Phi^{-1} \times \sum_s w_{h,s}[\cos\varphi(x+x_s,y+y_s)\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - \sin\varphi(x+x_s,y+y_s)\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}^*)]\mathbf{p}(x_s,y_s)$$
$$= \Phi^{-1}\sum_s w_{h,s}\sin(\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - \varphi(x+x_s,y+y_s))\mathbf{p}(x_s,y_s) \quad (46)$$

Inserting the last formula instead of $\partial_{\mathbf{c}}\partial_{\mathbf{c}^T}L_h(x,y,\mathbf{c}^*)$ in (40), we can solve this equation with respect to $\Delta\mathbf{c}$:

$$\Delta\mathbf{c} \simeq \Phi^{-1} \times \sum_s w_{h,s}[z_1(x+x_s,y+y_s)\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - z_2(x+x_s,y+y_s)\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}^*)]\mathbf{p}(x_s,y_s)$$
$$\Phi = \sum_s w_{h,s}\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s). \quad (44)$$

According to (28), the random estimation errors is

$$\Delta\mathbf{c}^0 \simeq \Phi^{-1} \times \sum_s w_{h,s}[\varepsilon_1(x+x_s,y+y_s)\sin\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - \varepsilon_2(x+x_s,y+y_s)\cos\tilde{\varphi}(x_s,y_s|\mathbf{c}^*)] \times \mathbf{p}(x_s,y_s)$$

where $\varepsilon_1 = -\sin\varphi(x+x_s,y+y_s)\cdot\Delta\varphi_s$, $\varepsilon_2 = \cos\varphi(x+x_s,y+y_s)\cdot\Delta\varphi_s$

Using these expressions for ε_1 and ε_2

$$\Delta\mathbf{c}^0 \simeq -\Phi^{-1} \times \sum_s w_{h,s}[\cos(\varphi(x+x_s,y+y_s)-\tilde{\varphi}(x_s,y_s|\mathbf{c}^*))]\mathbf{p}(x_s,y_s)\Delta\varphi_s.$$

If the estimates are accurate $\varphi \simeq \tilde{\varphi}$ and $\cos(\varphi-\tilde{\varphi}) \simeq 1$ and the covariance matrix of the random estimation errors $\Delta\mathbf{c}^0$ is calculated as

$$E\left\{\Delta\mathbf{c}^0\left(\Delta\mathbf{c}^0\right)^T\right\} = \frac{\sigma^2}{A^2}\Phi^{-1}\sum_s w_{h,s}^2\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s)\Phi^{-1} \quad (45)$$

where σ^2/A^2 is the variance of $\Delta\varphi_s$. For a symmetric window function, w_h, with $(x_s,y_s)\in U_h$, the polynomials $p_1=1$, $p_2=x_s$, $p_3=y_s$ are orthogonal on U_h and the matrices Φ and $\sum_s w_{h,s}^2\mathbf{p}(x_s,y_s)\mathbf{p}^T(x_s,y_s)$ are diagonal. Then the matrix $E\left\{\Delta\mathbf{c}^0\left(\Delta\mathbf{c}^0\right)^T\right\}$ is also diagonal. The first element of this matrix gives the formulas (35) for the estimate variance. Others give the variances of the derivative estimates.

For the bias evaluation, we consider the systematic part of (44) [see (46), shown at the top of the page].

Using (42), we have

$$|\sin(\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - \varphi(x+x_s,y+y_s))| \le |\tilde{\varphi}(x_s,y_s|\mathbf{c}^*) - \varphi(x+x_s,y+y_s)| \le L_2(x,y)\frac{(|x_s|+|y_s|)^2}{2}.$$

Then, $|E\{\Delta c_1|\} \le L_2(x,y)\sum_s w_{h,s}(|x_s|+|y_s|)^2/(2\sum_s w_{h,s})$. It proves (34) for the bias error of the estimates. Other items of the vector $E\{\Delta\mathbf{c}\}$ in (46) can be used in order to derive the bias of the derivative estimates.

ACKNOWLEDGMENT

The authors would like to thank the three anonymous reviewers for helpful and stimulating comments.

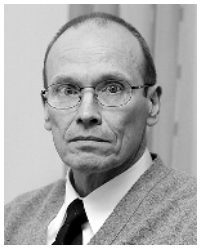
REFERENCES

[1] D. C. Ghiglia and M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*. New York: Wiley, 1998.
[2] T. Kreis, *Handbook of Holographic Interferometry (Optical and Digital Methods)*. Weinheim, Germany: Wiley, 2005.
[3] A. Patil and P. Rastogi, "Moving ahead with phase," *Opt. Lasers Eng.*, vol. 45, pp. 253–257, 2007.
[4] K. Itoh, "Analysis of the phase unwrapping algorithms," *Appl. Opt.*, vol. 21, pp. 2470–2470, 1982.
[5] J. Dias and J. Leitao, "The $Z\pi M$ algorithm: A method for interferometric image reconstruction in SAR/SAS," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 408–422, Apr. 2002.
[6] G. F. Carballo and P. W. Fieguth, "Member Hierarchical network flow phase unwrapping," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 8, pp. 1695–1708, Aug. 2002.
[7] G. Fornaro, A. Pauciullo, and E. Sansosti, "Phase difference-based multichannel phase unwrapping," *IEEE Trans. Image Process.*, vol. 14, no. 7, pp. 960–972, Jul. 2005.
[8] Q. Fang, P. M. Meaney, and K. D. Paulsen, "The multidimensional phase unwrapping integral and applications to microwave tomographical image reconstruction," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3311–3324, Nov. 2006.
[9] J. Dias and G. Valadão, "Phase unwrapping via graph cuts," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 684–697, Mar. 2007.
[10] S. Madsen, H. Zebker, and J. Martin, "Topographic mapping using radar interferometry: Processing techniques," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 1, pp. 246–256, Jan. 1993.
[11] A. Baldi, "Phase unwrapping by region growing," *Appl. Opt.*, vol. 42, no. 14, pp. 2498–2505, 2003.
[12] R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," *Radio Sci.*, vol. 23, pp. 713–720, 1988.
[13] M. Costantini, "A novel phase unwrapping method based on network programming," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 813–821, May 1998.
[14] D. L. Fried, "Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Amer.*, vol. 67, pp. 370–375, 1977.
[15] B. R. Hunt, "Matrix formulation of the reconstruction of phase values from phase differences," *J. Opt. Soc. Amer.*, vol. 69, pp. 393–399, 1979.
[16] D. C. Ghiglia and L. A. Romero, "Minimum L_p-norm two-dimensional phase unwrapping," *J. Opt. Soc. Amer. A*, vol. 13, pp. 1999–2013, 1996.
[17] 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.
[18] L. Ying, Z.-P. Liang, and D. C. Munson, Jr, "Unwrapping of MR phase images using a Markov random field model," *IEEE Trans. Med. Imag.*, vol. 25, no. 1, 2006.
[19] S. M. Song, S. Napel, N. J. Pelc, and G. H. Glover, "Phase unwrapping of MR phase images using Poisson equation," *IEEE Trans. Image Process.*, vol. 4, no. 5, pp. 667–676, May 1995.
[20] V. Katkovnik, K. Egiazarian, and J. Astola, *Local Approximation Techniques in Signal and Image Processing*. Bellingham, WA: SPIE, 2006.

[21] B. Friedlander and J. M. Francos, "An estimation algorithm for 2-D polynomial phase signals," *IEEE Trans. Image Process.*, vol. 5, no. 6, pp. 1084–1087, Jun. 1997.
[22] Z.-P. Liang, "A model-based method for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 893–897, Jun. 1996.
[23] V. Pascazio and G. Schirinzi, "Multifrequency insar height reconstruction through maximum likelihood estimation of local planes parameters," *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1478–1087, Dec. 2002.
[24] M. Servin, J. L. Marroquin, D. Malacara, and F. J. Cuevas, "Phase unwrapping with a regularized phase-tracking system," *Appl. Opt.*, vol. 37, no. 10, pp. 1917–1923, 1998.
[25] M. Servin, F. J. Cuevas, D. Malacara, J. L. Marroquin, and R. Rodriguez-Vera, "Phase unwrapping through demodulation by use of the regularized phase-tracking technique," *Appl. Opt.*, vol. 38, no. 10, pp. 1934–1941, 1999.
[26] D. M. Bates and D. G. Watts, *Nonlinear Regression Analysis and its Applications*. New York: Wiley, 1988.
[27] R. Fletcher, *Practical Methods of Optimization*. New York: Wiley, 1987.
[28] B. T. Polyak, *Introduction to Optimization*. New York: Optimization Software, 1987.
[29] C. W. Chen and H. A. Zebker, "Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms," *J. Opt. Soc. Amer. A*, vol. 17, no. 3, pp. 401–414, 2000.
[30] J. Leitao and M. Figueiredo, "Absolute phase image reconstruction: A stochastic non-linear filtering approach," *IEEE Trans. Image Process.*, vol. 7, no. 6, pp. 868–882, Jun. 1997.
[31] V. Katkovnik, J. Astola, and K. Egiazarian, "Noisy phase unwrap for holographic techniques: Adaptive local polynomial approximations," presented at the 3DTV Conf., Kos Island, Greece, 2007.

Vladimir Katkovnik received the M.Sc., Ph.D., and D.Sc. degrees in technical cybernetics from the Leningrad Polytechnic Institute, Leningrad, Russia, in 1960, 1964, and 1974, respectively.

From 1964 to 1991, he held the positions of Associate Professor and Professor at the Department of Mechanics and Control Processes, Leningrad Polytechnic Institute. From 1991 to 1999, he was a Professor of statistics with the Department of the University of South Africa, Pretoria. From 2001 to 2003, he was a Professor of mechatronics with the Kwangju Institute of Science and Technology, Korea. From 2000 to 2001, and since 2003, he has been a Research Professor with the Institute of Signal Processing, Tampere University of Technology, Tampere, Finland. He has published seven books and more than 200 papers. His research interests include stochastic signal processing, linear and nonlinear filtering, nonparametric estimation, imaging, nonstationary systems, and time-frequency analysis.



Jaakko Astola (F'00) received the Ph.D. degree in mathematics from Turku University, Finland, in 1978.

From 1976 to 1977, he was with the Research Institute for Mathematical Sciences, Kyoto University, Japan. Between 1979 and 1987, he was with the Department of Information Technology, Lappeenranta University of Technology, Finland. In 1984, he was a Visiting Scientist at the Eindhoven University of Technology, The Netherlands. From 1987 to 1992, he was an Associate Professor in applied mathematics at Tampere University, Finland. Since 1993, he has been a Professor of signal processing at Tampere University of Technology and is currently head of the Academy of Finland Centre of Excellence in Signal Processing. His research interests include signal processing, coding theory, spectral techniques, and statistics.

Karen Egiazarian (SM'96) was born in Yerevan, Armenia, in 1959. He received the M.Sc. degree in mathematics from Yerevan State University in 1981, the Ph.D. degree in physics and mathematics from Moscow State University, Moscow, Russia, in 1986, and the D.Tech. degree from the Tampere University of Technology (TUT), Tampere, Finland, in 1994.

He was a Senior Researcher with the Department of Digital Signal Processing, Institute of Information Problems and Automation, National Academy of Sciences of Armenia. Since 1996, he has been an Assistant Professor with the Institute of Signal Processing, Tampere University of Technology, where he is currently a Professor, leading the Transforms and Spectral Methods Group. His research interests are in the areas of applied mathematics, signal processing, and digital logic.