Minimum L^p-norm two-dimensional phase unwrapping

Dennis C. Ghiglia and Louis A. Romero

Sandia National Laboratories, Albuquerque, New Mexico 87185

Received January 25, 1996; revised manuscript received April 25, 1996; accepted May 17, 1996

We develop an algorithm for the minimum L^p-norm solution to the two-dimensional phase unwrapping problem. Rather than its being a mathematically intractable problem, we show that the governing equations are equivalent to those that describe weighted least-squares phase unwrapping. The only exception is that the weights are data dependent. In addition, we show that the minimum L^p-norm solution is obtained by embedding the transform-based methods for unweighted and weighted least squares within a simple iterative structure. The data-dependent weights are generated within the algorithm and need not be supplied explicitly by the user. Interesting and useful solutions to many phase unwrapping problems can be obtained when $p < 2$. Specifically, the minimum L^0-norm solution requires the solution phase gradients to equal the input data phase gradients in as many places as possible. This concept provides an interesting link to branch-cut unwrapping methods, where none existed previously.

1. INTRODUCTION

Two-dimensional (2-D) phase unwrapping continues to capture the interest and the imagination of researchers in various disciplines. It is interesting that so many diverse applications share the common requirement to obtain some useful form of unwrapped phase from samples of wrapped values on a rectangular grid. A few of these applications are synthetic aperture radar (SAR) interferometry for terrain elevation estimation,[1–7] optical interferometry for surface characterization,[8,9] wave-front compensation for imaging through aberrating media,[10–12] and magnetic resonance imaging (MRI) for the water/fat separation problem.[13–15]

In the most general and least restrictive sense 2-D phase unwrapping is an impossible problem. For example, an unknown phase function $\varphi(x, y)$ corrupted by noise and wrapped into the interval $(-\pi, \pi)$ is impossible to recover unambiguously. Impossible though it may be, this does not mean that we should give up! On the contrary, certain assumptions of the underlying process or the nature of the desired solution makes the phase unwrapping problem tractable, if not easy. More importantly, many useful solutions are obtained from rather simple assumptions.

The most common notion is to assume that the desired unwrapped phase has local phase gradients (differences) that are less than π rad in magnitude everywhere. Specifically, we assume that we know the phase, ϕ mod 2π, of a function on a discrete grid of points:

$$\psi_{i,j} = \phi_{i,j} + 2\pi k_{i,j}, \qquad k_{i,j} \text{ an integer array,}$$
$$-\pi < \psi_{i,j} \leq \pi, \quad i = 0, \ldots, M-1,$$
$$j = 0, \ldots, N-1. \tag{1}$$

Given the wrapped phase values $\psi_{i,j}$, we wish to determine the unwrapped phase values $\phi_{i,j}$ at the same grid locations. The requirement is that the phase differences of the $\phi_{i,j}$ be less than π rad in magnitude everywhere. In this case ϕ includes the effect of noise on the unknown phase φ. We are attempting only to obtain ϕ, which hopefully captures some useful embodiment of the unknown and unobtainable φ. A common path-following methodology for the solution is to start at some arbitrary grid point sample ψ_{i_0,j_0} and to add some multiple of 2π to an adjacent sample so that the absolute value of the spatial phase difference is less than π rad in magnitude. A path is selected that covers the entire array. An alternative, but equivalent, methodology for path-following unwrapping is to begin at some arbitrary grid point sample and integrate wrapped spatial phase differences over a path that covers the entire array.[16]

Wrapped phase differences (with respect to the i and j indices) are defined as

$$f_{i,j} = \begin{cases} W\{\psi_{i+1,j} - \psi_{i,j}\} & i = 0, \ldots, M-2, \\ & j = 0, \ldots, N-1, \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$g_{i,j} = \begin{cases} W\{\psi_{i,j+1} - \psi_{i,j}\} & i = 0, \ldots, M-1, \\ & j = 0, \ldots, N-2. \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The wrapping operator, W, wraps all values of its argument into the range $(-\pi, \pi)$ by adding or subtracting an integral number of 2π rad from its argument.

Unfortunately, it is not possible to unwrap the data unambiguously and meet the local gradient assumption everywhere if the wrapped phase data contain residues. Residues exist whenever the sum of wrapped phase differences around a closed path differs from zero.[2,17] For example, we compute

$$r_{i,j} = g_{i,j} + f_{i,j+1} - g_{i+1,j} + f_{i,j},$$
$$i = 0, \ldots, M-1, \qquad j = 0, \ldots, N-1. \tag{4}$$

If $r_{i,j} > 0$, we have a positive residue at sample (i, j). If $r_{i,j} < 0$, we have a negative residue at sample (i, j). Otherwise, if $r_{i,j} = 0$, no residue exists. Equation (4)

0740-3232/96/1001999-15$10.00

represents a path integral around a closed 2×2 sample box beginning and ending with sample (i, j). If no residues are found in the entire $M \times N$ array, the wrapped phase $\psi_{i,j}$ can be uniquely unwrapped (to within an arbitrary constant) by any unwrapping procedure.

The fun begins, of course, when ψ contains residues. Path-following schemes will generate path-dependent results unless cut lines or branch cuts are inserted between residues of opposite polarity, from residues to the array boundary, or combinations thereof. These cut lines act as barriers to the unwrapping by preventing the unwrapping path from crossing any cut lines or encircling unpaired opposite-polarity residues.[2,18] The spatial phase differences (of the unwrapped phase) can then be larger than π rad in magnitude across the cut lines. Considerable effort has gone into developing residue paring and cut line generation methods.[19–23] Because ψ provides residue polarities and locations only, a great many cut lines are possible. Many heuristic and *ad hoc* rules are imposed to create cut lines that yield, to some degree, an unwrapped phase having certain other useful properties. (These properties are based, in part, on the researcher's biases, preferences, or other not easily imposed assumptions.) In other words, the unwrapped phase ϕ appears to embody many physical, useful, or other important properties of the true underlying phase φ. Path-following methods can, in many applications, provide very good results. They should be part of a well-stocked arsenal of phase unwrapping methods.

Another approach to 2-D phase unwrapping imposes constraints on the solution in a mathematically formal manner. For example, given the wrapped phase values $\psi_{i,j}$, we wish to determine the unwrapped phase values $\phi_{i,j}$ at the same grid locations with the requirement that the phase differences of the $\phi_{i,j}$ agree with those of the $\psi_{i,j}$ in the least-squares sense. The solution, $\phi_{i,j}$, that minimizes

$$\varepsilon^2 = \sum_{i=0}^{M-2} \sum_{j=0}^{N-1} (\phi_{i+1,j} - \phi_{i,j} - f_{i,j})^2 + \sum_{i=0}^{M-1} \sum_{j=0}^{N-2} (\phi_{i,j+1} - \phi_{i,j} - g_{i,j})^2 \qquad (5)$$

is the least-squares solution. Least-squares phase unwrapping problems, formulated in both unweighted and weighted forms, amount to solving discretized partial differential equations (PDE's) with efficient mathematical methods.[24–29] Additional constraints, such as smoothness, can be imposed through regularization methods discussed specifically in Ref. 30 and more generally in Ref. 31. In general, many L^2-norm minimization problems are posed primarily because the resulting mathematics is tractable and/or the solution is amenable to efficient methods.

Least-squares methods, unlike path-following methods, do not deal directly with the residue problem because the solution is obtained by integrating over the residues to minimize gradient differences, as shown in Eq. (5) (for the unweighted case). Weighted least squares requires the explicit definition of weights to attack any shortcomings of unweighted least squares. Specifically, weights are selected to accommodate residues in some fashion, to isolate regions of low signal-to-noise ratio, or to impose other ill-defined properties or preferences on the resultant solution. In some applications the unweighted least-squares solution may be preferred to a path-following result because of the global smoothness constraint. In other cases weighted or unweighted least-squares solutions that have not adequately accommodated the residue problem can provide undesirable results. Any uncompensated residues produce least-squares solutions with underestimated (in magnitude) phase gradients everywhere.[32,33] The severity of this problem depends on the density of the uncompensated residues and on the specific application at hand.

A colleague, Mark Pritt, of Loral Federal Systems, Gaithersburg, Md., suggested that a useful unwrapped phase would result if, instead of the least-squares solution (L^2-norm minimization), the minimum L^1-norm solution were easily computable.[34] It is easy to dismiss this concept at first glance because of the possibility of intractable mathematics and/or difficult computations. Surprisingly, that is not the case at all. In fact, it is relatively straightforward to obtain not just the minimum L^1-norm solution but the more general minimum L^p-norm solution. In addition, the computational approach builds on and uses the means of solving the weighted least-squares phase unwrapping problem provided as Algorithm 3 in Ref. 29. In contrast to weighted least squares, for which explicit weights must be provided, L^p-norm algorithms effectively generate their own gradient weights from the input data when $p \neq 2$. Some interesting and potentially valuable solutions to the 2-D phase unwrapping problem can be obtained from this approach. Several examples will be provided in Section 4.

We believe that while there is no best unwrapping algorithm, there may be a best approach for a particular application. Therefore, in the spirit of fostering further research, we provide a minimum L^p-norm 2-D phase unwrapping algorithm.

2. MATHEMATICAL DEVELOPMENT

Given the wrapped phase values $\psi_{i,j}$, we wish to determine the unwrapped phase values $\phi_{i,j}$ at the same grid locations. We require that the phase differences of the $\phi_{i,j}$ agree with those of the $\psi_{i,j}$ in the minimum L^p-norm sense. That is, the solution, $\phi_{i,j}$, that minimizes

$$J = |\varepsilon|^p = \sum_{i=0}^{M-2} \sum_{j=0}^{N-1} |\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^p + \sum_{i=0}^{M-1} \sum_{j=0}^{N-2} |\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^p \qquad (6)$$

is the minimum L^p-norm solution. Interesting possibilities occur when $p \neq 2$, as we will see.

A. Continuous Variational Approach

We will take two slightly different approaches in developing the solution to the above minimization problem. Each approach provides a slightly different insight. The

first approach rewrites Eq. (6) in integral form as if we were dealing with continuous functions. Thus we wish to find the function $\phi(x, y)$ that minimizes

$$J = \int\int f(\phi_x, \phi_y, x, y)\mathrm{d}x\mathrm{d}y, \tag{7}$$

where

$$f = |\phi_x - \psi_x|^p + |\phi_y - \psi_y|^p \tag{8}$$

and phase differences have been replaced with corresponding partial derivatives. (Note that ϕ_x and ϕ_y refer to partial derivatives and should not be confused with the subscripted indices $\phi_{i,j}$; the same holds analogously for ψ_x and ψ_y.) Calculus of variations is used to find the conditions under which J assumes a stationary value. The function ϕ minimizing J must therefore be the solution of the PDE resulting from the appropriate Euler–Lagrange equation (see Refs. 35 and 36):

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial \phi_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial \phi_y}\right) = 0. \tag{9}$$

We use the identity

$$\frac{\partial}{\partial z}(|z|^p) = \frac{\partial}{\partial z}[(z^2)^{p/2}] = pz|z|^{p-2} \tag{10}$$

to compute the partial derivatives of f [Eq. (8)] with respect to ϕ_x and ϕ_y and obtain

$$\frac{\partial f}{\partial \phi_x} = p(\phi_x - \psi_x)|\phi_x - \psi_x|^{p-2},$$

$$\frac{\partial f}{\partial \phi_y} = p(\phi_y - \psi_y)|\phi_y - \psi_y|^{p-2}. \tag{11}$$

Substituting Eq. (11) into Eq. (9) yields the nonlinear PDE that we must solve to obtain the minimum L^p-norm solution for ϕ:

$$\left\{\frac{\partial}{\partial x}[(\phi_x - \psi_x)|\phi_x - \psi_x|^{p-2}] + \frac{\partial}{\partial_y}[(\phi_y - \psi_y)|\phi_y - \psi_y|^{p-2}]\right\} = 0. \tag{12}$$

With the substitutions

$$U(x, y) = |\phi_x - \psi_x|^{p-2}, \qquad V(x, y) = |\phi_y - \psi_y|^{p-2} \tag{13}$$

Eq. (12) becomes

$$\left\{\frac{\partial}{\partial x}[U(x, y)(\phi_x - \psi_x)] + \frac{\partial}{\partial y}[V(x, y)(\phi_y - \psi_y)]\right\} = 0. \tag{14}$$

Note that in the special case in which $U(x, y) = V(x, y) = 1$ we obtain Poisson's equation

$$\phi_{xx} + \phi_{yy} = \psi_{xx} + \psi_{yy}, \tag{15}$$

or

$$\nabla^2\phi = \rho \tag{16}$$

as expected.[29] [Notice that $U(x, y) = V(x, y) = 1$ when $p = 2$.]

Now let us assume for the moment that both $U(x, y)$ and $V(x, y)$ are generalized weighting functions [which weight the phase gradients as shown in Eq. (14)] independent of the solution ϕ. With this assumption Eq. (14) is equivalent to the characteristic equation describing weighted least-squares phase unwrapping. An efficient algorithm for solving the weighted least-squares phase unwrapping problem is described in Ref. 29.

It can now be seen that if we stay away from least-squares problems, the generalized weights, $U(x, y)$ and $V(x, y)$ (for $p \neq 2$), are functions of both the original data ψ and the desired solution ϕ as given in Eqs. (13). Therefore we begin to see the basis for an iterative algorithm for solving Eq. (14) in the general case. Specifically, $U_0(x, y)$ and $V_0(x, y)$ would be computed from the original data ψ and an initial guess for ϕ_0. These initial weights are held fixed in Eq. (14), which is then solved by Algorithm 3 given in Ref. 29 (or other efficient weighted least-squares solvers) (see Refs. 37 and 38). The current solution ϕ_1 and the original data ψ are then used to compute a new set of weights, $U_1(x, y)$ and $V_1(x, y)$, from which the next estimate ϕ_2 is obtained. This process is repeated until convergence is achieved. Although the previous discussion has provided the general notion for a workable algorithm, many subtleties need to be addressed. We discuss these and other issues in later sections.

B. Discrete Variational Approach

Let us now look at the variational approach starting with the actual discretized minimization problem embodied in Eq. (6). We will include a fair amount of detail in this approach because it forms the basis of our working algorithm.

The total variation of J, denoted by δJ, consists of the contributions of the variations depending on ϕ. With reference to Eq. (6) [and use of Eq. (10)] we obtain

$$\delta J = p(\delta_1 + \delta_2). \tag{17}$$

where

$$\delta_1 = \sum_{i=0}^{M-2}\sum_{j=0}^{N-1}(\phi_{i+1,j} - \phi_{i,j} - f_{i,j})|\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^{p-2}(\delta\phi_{i+1,j} - \delta\phi_{i,j}), \tag{18}$$

$$\delta_2 = \sum_{i=0}^{M-1}\sum_{j=0}^{N-2}(\phi_{i,j+1} - \phi_{i,j} - g_{i,j})|\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^{p-2}(\delta\phi_{i,j+1} - \delta\phi_{i,j}). \tag{19}$$

We now introduce the following equations for economy of notation:

$$a_{i,j} = (\phi_{i+1,j} - \phi_{i,j} - f_{i,j})|\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^{p-2}, \qquad 0 \leq i \leq M-2, \qquad 0 \leq j \leq N-1, \tag{20}$$

with the boundary conditions

$$a_{i,j} = 0, \qquad i = -1, \qquad i = M-1, \qquad 0 \leq j \leq N-1. \tag{21}$$

$$b_{i,j} = (\phi_{i,j+1} - \phi_{i,j} - g_{i,j})|\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^{p-2}, \qquad 0 \leq i \leq M-1, \qquad 0 \leq j \leq N-2, \tag{22}$$

with the boundary conditions

$$b_{i,j} = 0, \qquad j = -1, \qquad j = N-1, \qquad 0 \leqslant i \leqslant M-1. \tag{23}$$

Substitution of Eqs. (20) and (22) into Eqs. (18) and (19), respectively, yields

$$\delta_1 = \sum_{i=0}^{M-2} \sum_{j=0}^{N-1} a_{i,j}(\delta\phi_{i+1,j} - \delta\phi_{i,j}), \tag{24}$$

$$\delta_2 = \sum_{i=0}^{M-1} \sum_{j=0}^{N-2} b_{i,j}(\delta\phi_{i,j+1} - \delta\phi_{i,j}). \tag{25}$$

Now, in order to sum the above contributions and set $\delta J = 0$, we need to shift the i and the j indices of Eqs. (24) and (25), respectively, so that we have a common variation $\delta\phi_{i,j}$ multiplying the individual contributions. Rearranging the summations as required yields the following equations (note the change in summation limits):

$$\delta_1 = -\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (a_{i,j} - a_{i-1,j})\delta\phi_{i,j}, \tag{26}$$

$$\delta_2 = -\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (b_{i,j} - b_{i,j-1})\delta\phi_{i,j}. \tag{27}$$

Note that the above equations assume that $a_{i,j} = b_{i,j} = 0$ on the boundary, as stated in Eqs. (21) and (23).

Substitution of Eqs. (26) and (27) into Eq. (17) yields

$$\delta J = -p \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (b_{i,j} - b_{i,j-1} + a_{i,j} - a_{i-1,j})\delta\phi_{i,j}. \tag{28}$$

Assuming that the perturbation $\delta\phi_{i,j}$ is arbitrary, we conclude that

$$b_{i,j} - b_{i,j-1} + a_{i,j} - a_{i-1,j} = 0, \qquad 0 \leqslant i \leqslant M-1, \ 0 \leqslant j \leqslant N-1. \tag{29}$$

Equation (29) requires that we know $\phi_{-1,j}$, $\phi_{M,j}$, $\phi_{i,-1}$, and $\phi_{i,N}$ and the corresponding values of ψ. However, Eqs. (21) and (23) can be thought of as boundary conditions that determine these values in terms of the interior values. Thus

$$\begin{aligned} \phi_{-1,j} - \phi_{0,j} &= \psi_{-1,j} - \psi_{0,j}, \qquad 0 \leqslant j \leqslant N-1, \\ \phi_{M,j} - \phi_{M-1,j} &= \psi_{M,j} - \psi_{M-1,j}, \qquad 0 \leqslant j \leqslant N-1, \\ \phi_{i,-1} - \phi_{i,0} &= \psi_{i,-1} - \psi_{i,0}, \qquad 0 \leqslant i \leqslant M-1, \\ \phi_{i,N} - \phi_{i,N-1} &= \psi_{i,N} - \psi_{i,N-1}, \qquad 0 \leqslant i \leqslant M-1. \end{aligned} \tag{30}$$

Finally, expanding Eq. (29) and imposing the boundary conditions of Eqs. (30) yield the following set of equations, whose solution represents the minimum L^p norm:

$$\begin{aligned} &(\phi_{i+1,j} - \phi_{i,j} - f_{i,j})U(i,j) + (\phi_{i,j+1} - \phi_{i,j} - g_{i,j}) \\ &\quad \times V(i,j) - (\phi_{i,j} - \phi_{i-1,j} - f_{i-1,j})U(i-1,j) \\ &\quad - (\phi_{i,j} - \phi_{i,j-1} - g_{i,j-1})V(i,j-1) = 0, \end{aligned} \tag{31}$$

where

$$U(i,j) = \begin{cases} |\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^{p-2} & i = 0, \ldots, M-2 \\ & j = 0, \ldots, N-1, \\ 0 & \text{otherwise} \end{cases} \tag{32}$$

$$V(i,j) = \begin{cases} |\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^{p-2} & i = 0, \ldots, M-1 \\ & j = 0, \ldots, N=2. \\ 0 & \text{otherwise} \end{cases} \tag{33}$$

The rather formidable-looking results shown in Eqs. (31)–(33) are just the discrete counterpart to Eq. (14) with the boundary conditions imposed and the indices properly aligned.

We see that when $p = 2$, Eq. (31) reduces to the discrete version of Poisson's equation,

$$(\phi_{i+1,j} - 2\phi_{i,j} + \phi_{i-1,j}) + (\phi_{i,j+1} - 2\phi_{i,j} + \phi_{i,j-1}) = \rho_{i,j}, \tag{34}$$

where

$$\rho_{i,j} = f_{i,j} - f_{i-1,j} + g_{i,j} - g_{i,j-1}. \tag{35}$$

Equation (31) can now be solved by exploiting the efficient weighted least-squares unwrapping method (Algorithm 3) of Ref. 29 by imbedding it within an outer loop. The loop's sole purpose is to update the data-dependent weights $U(i,j)$ and $V(i,j)$. [See Appendix A for a review of the unweighted and weighted least-squares algorithms (Algorithms LS and WLS, respectively).]

C. Minimum L^p-Norm Two-Dimensional Phase Unwrapping

We have shown that the minimum L^p-norm solution requires solving a nonlinear PDE [see Eq. (14) (continuous version) and Eq. (31) (discrete version)]. It is nonlinear because the weights, U and V, are functions of the input data and of the solution. We have hinted at an iterative scheme to solve this problem. We now make the structure of the algorithm explicit. However, it should be noted that there are other practical issues that need to be addressed before the algorithm can be considered complete. We will address some of these practical issues in Section 3, but for now we present the simplified algorithm.

Let us begin by rearranging Eq. (31) into the desired form as follows:

$$\begin{aligned} &(\phi_{i+1,j} - \phi_{i,j})U(i,j) + (\phi_{i,j+1} - \phi_{i,j})V(i,j) - (\phi_{i,j} \\ &\quad - \phi_{i-1,j})U(i-1,j) - (\phi_{i,j} - \phi_{i,j-1})V(i,j-1) \\ &\quad = c(i,j), \end{aligned} \tag{36}$$

where

$$c(i,j) = f_{i,j}U(i,j) - f_{i-1,j}U(i-1,j) + g_{i,j}V(i,j) - g_{i,j-1}V(i,j-1), \tag{37}$$

$$U(i,j) = \begin{cases} |\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^{p-2} & i = 0,\ldots,M-2,\ j = 0,\ldots,N-1 \\ 0 & \text{otherwise} \end{cases}, \tag{38}$$

$$V(i,j) = \begin{cases} |\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^{p-2} & i = 0,\ldots,M-1,\ j = 0,\ldots,N-2 \\ 0 & \text{otherwise} \end{cases}. \tag{39}$$

We represent Eq. (36) in matrix form as

$$\mathbf{Q}\phi = \mathbf{c}. \tag{40}$$

The simplified algorithm for solving the minimum L^p-norm 2-D phase unwrapping problem is given below.

Algorithm LP-NORM-S (simplified version; provides structure only and is not used in practice):

1. $l = 0$, $\phi_l = 0$.
2. Compute data-dependent weights, $U_l(i,j)$ and $V_l(i,j)$, from Eqs. (38) and (39).
3. Compute $c_l(i,j)$ from Eq. (37) and use $U_l(i,j)$ and $V_l(i,j)$ from step 2.
4. Solve Eq. (40), $\mathbf{Q}\phi_l = \mathbf{c}_l$, with the use of Algorithm WLS.
5. $l = l + 1$.
6. If $l \geqslant l_{\max}$, end; otherwise go to step 2.

Although the above algorithm highlights the correct iterative structure, it does not work well in practice for several reasons. We will now address several modifications to Algorithm LP-NORM-S that do indeed make a working algorithm.

3. IMPLEMENTATION, CONNECTION TO BRANCH-CUT METHODS, AND OTHER PRACTICAL ISSUES

Because we are free to select the norm any way we want, what does it mean to select a specific value for p? We know that for large p (i.e., $p \gg 2$) solution gradients that are far from the corresponding measured gradients are penalized very severely. All remaining solution gradients are adjusted so that the outliers have minimum deviation. Therefore norms for which $p \gg 2$ provide solutions that tend to be very smooth globally, and outliers strongly pull the solution in their direction. This is not desired for 2-D phase unwrapping because the solution phase gradients will not equal the data gradients anywhere. Besides, $p \gg 2$ exaggerates all the undesirable properties of least-squares solutions.

On the other hand, when $p < 2$, the solution tends to have local gradients that exactly match the measured gradients in as many places as possible. We can see this effect through the data-dependent weights of Eqs. (38) and (39). Specifically, when $p < 2$, the weights become large when the gradients of the current solution are close to the measured gradients. We can make this notion precise when $p \equiv 0$. In this case Eq. (6) gives a value for J equal to the sum of the number of samples for which the solution gradients do not exactly equal the measured gradients. This, it appears, is exactly what we desire. In other words, the minimum L^0-norm solution is such that the solution gradients exactly equal the measured gradients in as many places as possible. By comparison, if one unwraps a phase map after placing branch cuts, the resulting phase gradients are identical to those of the measured data except across the branch cuts. These notions give credence to the desire to minimize the sum of cut lengths as described in Ref. 39. Thus the minimum L^0-norm solution provides an interesting and thought-provoking link to branch-cut methods where none existed previously.

For $p \equiv 0$ the solution gradients deviate from the measured gradients in as few places as possible; thus the magnitude of the deviation is of no concern (when $p \equiv 0$). The magnitude of the deviation is of concern for $p \neq 0$, however. There appears to be a smooth transition of solution properties as p varies. Thus, for example, the $p = 1.9$ solution should be quite close to the $p = 2$ (unweighted least-squares) solution. We are concerned only with minimum L^p-norm solutions for $p < 2$. In practice, we will try to obtain only the $p = 0$ solution for the following reasons: (1) The $p = 0$ solution has the optimality properties and the link to branch-cut methods as described above. (2) While $p = 1$ solutions have been obtained, they appear to provide no particular advantage. In fact, we have observed slower and sometimes erratic convergence behavior for $p = 1$. This may be due in part to algorithm-tuning issues referred to below. It appears that $p = 0$ computations are more well behaved, although we have not extensively explored these solution variations.

In theory, any values for weights are acceptable. In practice, weighted (e.g., weighted least-squares) problems are better behaved numerically when the weights are normalized to lie in the range (0, 1). The problem is to obtain some useful form of normalization of the weights given in Eqs. (38) and (39). We have found that the following equations provide a means for normalizing the data-dependent weights:

$$U(i,j) = \frac{\varepsilon_0}{|\phi_{i+1,j} - \phi_{i,j} - f_{i,j}|^{2-p} + \varepsilon_0}, \tag{41}$$

$$V(i,j) = \frac{\varepsilon_0}{|\phi_{i,j+1} - \phi_{i,j} - g_{i,j}|^{2-p} + \varepsilon_0}. \tag{42}$$

We have found that setting $\varepsilon_0 = 0.01$ (not to be confused with ε in step 7 of Algorithm WLS; see Appendix A) with $p = 0$ provides a reasonable compromise between how closely we require the solution gradients to match the measured gradients and the rate of convergence of Algorithm WLS. For example, if ε_0 is too small, Algorithm WLS is slow to converge (i.e., too many inner iterations, of index k, are required). In addition, too many outer iterations (i.e., index l) in Algorithm LP-NORM-S (for example) may be required. We will discuss typical num-

bers for inner and outer iterations in Section 4. There is another modification that can be made to speed convergence to a useful minimum L^p-norm solution, which we present in Subsection 3.A.

A. Unwrapping the Residual

We mentioned in Section 1 that any wrapped phase is uniquely unwrappable (to within an arbitrary constant) if it does not contain residues. This is true for the wrapped residual between the input wrapped phase and any partially converged estimate $\phi_l(i, j)$. Thus, at the beginning of each outer iteration (iteration l), we compute a residual wrapped phase

$$R(i, j) = W\{\psi(i, j) - \phi_l(i, j)\}. \tag{43}$$

Next we test $R(i, j)$ for residues by using the procedure outlined in Eq. (4) and in the accompanying text. Note that the wrapped phase differences in Eq. (4) [with Eqs. (2) and (3)] come from $R(i, j)$ and not $\psi(i, j)$ in this case. If no residues exist, $R(i, j)$ can be unwrapped uniquely with Algorithm LS (or, for that matter, a simple path-following method). The unwrapped residual is added back to $\phi_l(i, j)$ to form the final estimate

$$\phi(i, j) = \phi_l(i, j) + W^{-1}\{R(i, j)\}, \tag{44}$$

where W^{-1} represents a unique unwrapping operation. Unwrapping the residual accomplishes two things: (1) It ensures that the final unwrapped phase differs from the measured values by only integral multiples of 2π, and (2) it speeds up the time to final solution by not having to wait for complete convergence from the algorithm outer loop. In practice, as long as the weighted least-squares solution from step 8 below is sufficiently converged, we always seem to obtain a residue-free residual at some point. In other words, if the algorithm will indeed converge to a minimum L^p-norm solution, the residual obviously will become residue free (the residual goes toward zero).

B. The Minimum L^p-Norm Algorithm

We now incorporate all these concepts into the following minimum L^p-norm 2-D phase unwrapping algorithm.

Algorithm LP-NORM (minimum L^p-norm 2-D phase unwrapping algorithm):

1. Select $p < 2$. Suggest that $p = 0$.
2. Initialize outer iteration index and initial solution ($l = 0$, $\phi_l = 0$).
3. Compute residual $R(i, j)$ from Eq. (43).
4. Test $R(i, j)$ for residues. If no residues exist, go to step 5; otherwise, go to step 6.
5. Unwrap $R(i, j)$ (with Algorithm LS or simple path-following method), and correct current estimate with Eq. (44). Compute final normalized data-dependent weights, $U_l(i, j)$ and $V_l(i, j)$, from Eqs. (41) and (42). Done.
6. Compute normalized data-dependent weights, $U_l(i, j)$ and $V_l(i, j)$, from Eqs. (41) and (42).
7. Compute $c_l(i, j)$ from Eq. (37) and use $U_l(i, j)$ and $V_l(i, j)$ from step 6.
8. Solve Eq. (40), $\mathbf{Q}\phi_l = \mathbf{c}_l$, with Algorithm WLS.
9. Increment outer iteration index ($l = l + 1$).
10. If $l \geq l_{\max}$, end; otherwise, go to step 3.

Note that we compute, one last time, the data-dependent weights from the converged solution as indicated in step 5. This allows us to view the final spatial distribution of the weights to help assess the quality and the character of the solution. In some of the examples to follow in Section 4 we show the progression of the weights as a function of the outer iteration index l.

One of the reviewers pointed out that the ability of Algorithm 3, given in Ref. 29, or of Algorithm WLS, given in Appendix A, to interpolate smoothly through zero-weighted regions is a crucial element of this new algorithm. We would have to agree. After all, if a weighted least-squares solution is unavailable (or "garbage") in regions of low-valued weights, this "bad" solution would provide no assistance in obtaining a better set of weights for the next iteration. The process could stagnate at best or diverge at worst.

4. EXAMPLES OF TWO-DIMENSIONAL PHASE UNWRAPPING

All images depicting wrapped phases in the range $(-\pi, \pi)$ are scaled between black and white (i.e., 0–255) for display. Unwrapped phase images are also scaled between black and white to capture the full dynamic range.

Let us begin our examples with the wrapped values of a 128×128 pixel 2-D phase surface with shear as shown in Fig. 1(a). These data cannot be uniquely unwrapped because they contain five negative residues located along the perceived shear line, as shown in Fig. 1(b). The residue locations are shown as black dots on a uniform gray background. The origin of these figures, and of all other figures in this paper, is indicated in Fig. 1(a).

Our intuition (or preference) indicates that a horizontal shear line exists midway between the top and the bottom of Fig. 1(a). Thus we expect the correct unwrapped phase to tear along this shear line and nowhere else. This is such a strong notion that any other tear location would be considered incorrect or undesirable. It is interesting to watch the evolution of the data-dependent weights. Also note the unwrapped phase at each iteration toward the minimum L^p-norm solution ($p = 0$ in this and all other examples). We used 30 iterations (k index) in Algorithm WLS for this example. We often re-

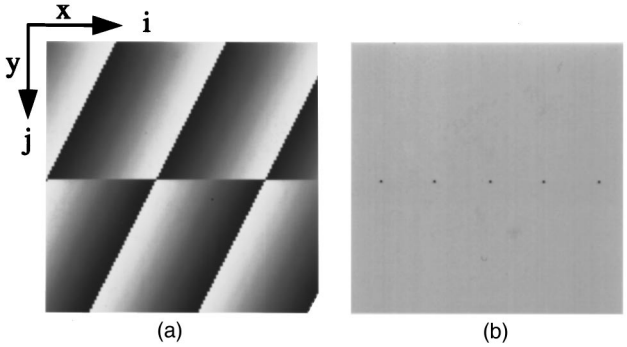

Fig. 1. (a) Wrapped values of a 128×128 pixel 2-D phase surface with shear (scaled for display). (b) The five negative residues are located at the black dots. No other residues are present.

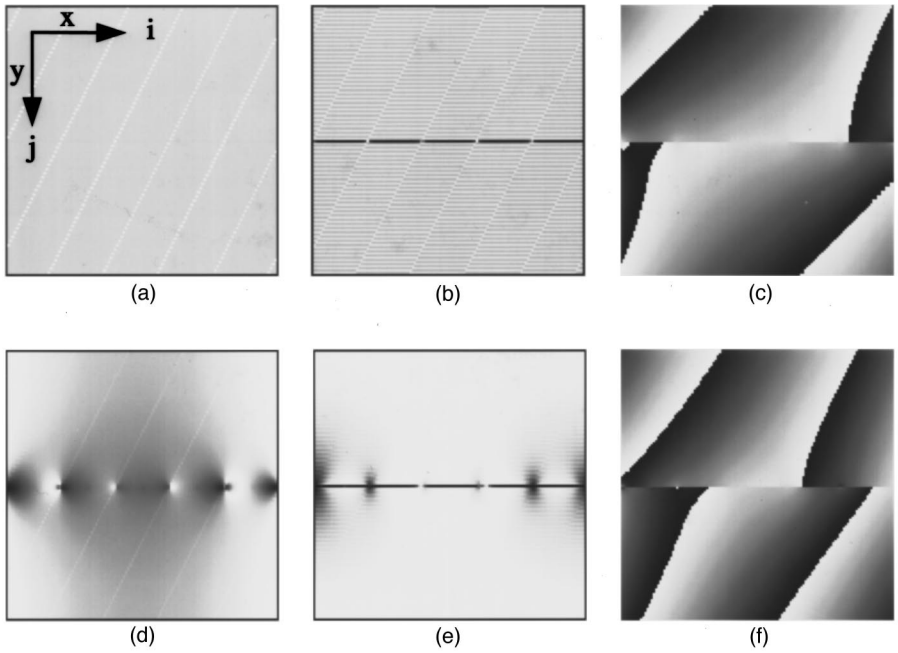

Fig. 2. Progress of Algorithm LP-NORM on the phase-shear problem. In this and all other examples in this paper we set $p = 0$. (a) Data-dependent x-gradient weights, $U_0(i, j)$, from the initial guess, $\phi(i, j) = 0$; (b) data-dependent y-gradient weights, $V_0(i, j)$, from the initial guess, $\phi(i, j) = 0$; rewrapped solution after iteration $l = 1$; (d) $U_1(i, j)$; (e) $V_1(i, j)$; (f) rewrapped solution after iteration $l = 2$.

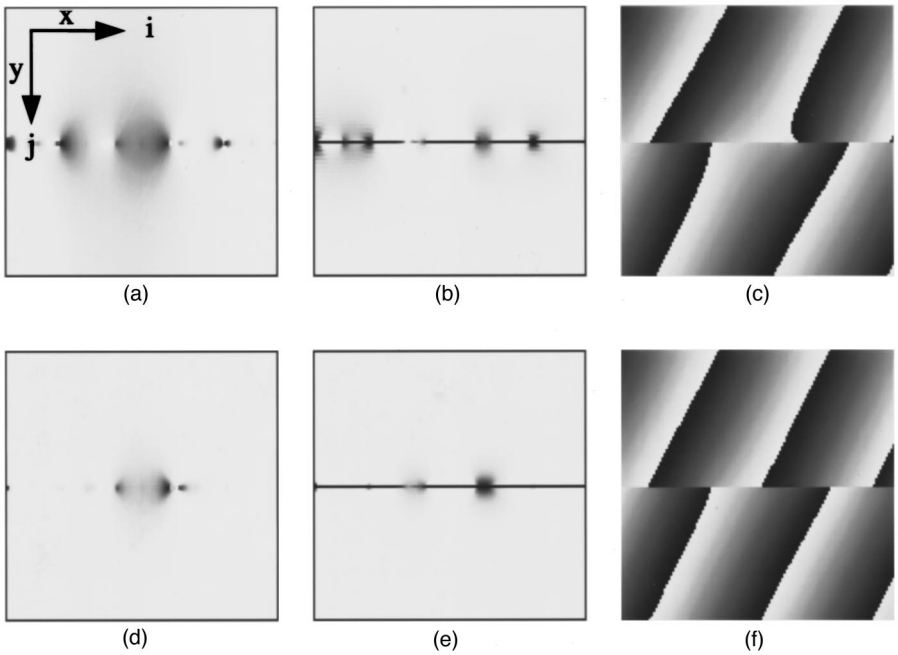

Fig. 3. Continued progress of Algorithm LP-NORM on the phase-shear problem: (a) $U_2(i, j)$, (b) $V_2(i, j)$, (c) rewrapped solution after iteration $l = 3$, (d) $U_3(i, j)$, (e) $V_3(i, j)$, (f) rewrapped solution after iteration $l = 4$.

wrap the unwrapped result to show more clearly the progress of the algorithm.

Figures 2 and 3 show the progress of Algorithm LP-NORM on the data depicted in Fig. 1(a). Note that a 1-pixel-wide black border is drawn around the data-dependent weight figures. The border simply outlines the array boundaries and is not to be interpreted as part of the weights.

Convergence on the shear problem was achieved after five outer iterations. The unwrapped phase, shown in Fig. 4(f), tears along the perceived shear line, as our intuition demanded. More importantly though, for this example, the unwrapped phase is indeed the minimum L^0-norm solution. Analysis of the y-gradient weights $V(i, j)$, shown in Fig. 4(e), indicates that there are only 101 samples at which the gradients of the unwrapped

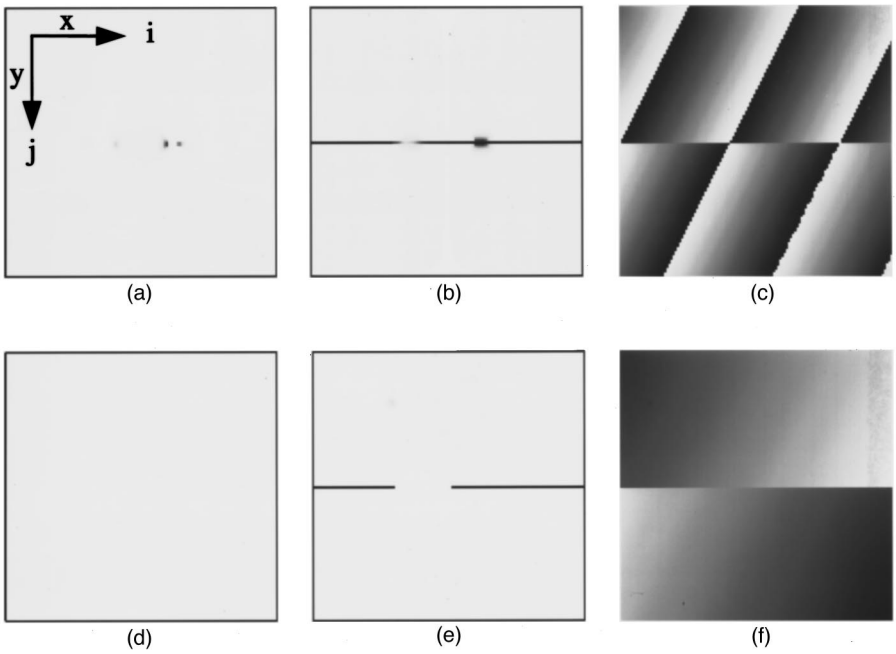

Fig. 4. Convergence on the phase-shear problem. (a) $U_4(i, j)$. (b) $V_4(i, j)$. (c) Rewrapped solution after iteration $l = 5$. This is the result obtained (at convergence) after correcting the solution with the unwrapped residual. It agrees perfectly with the original input, as it should. (d) The x gradient weights, $U_5(i, j)$, are all unity at convergence because there is no disagreement between the x gradients of the solution and the x gradients of the input data. (e) The y-gradient weights, $V_5(i, j)$, differ from unity at only 101 locations. Based on the locations of the residues, this is the minimum number of locations at which the gradients of the solution can differ from the gradients of the original data. For this example, a minimum L^0-norm solution was obtained. (f) Unwrapped minimum L^0-norm solution after iteration $l = 5$.

phase do not exactly equal the gradients of the original wrapped input data. With knowledge of the exact locations of the residues [Fig. 1(b)], we determined that there are no other ways in which fewer gradient weights could be selected to yield an unwrapped result. In other words, the unwrapped phase has gradients that correspond exactly with the gradients of the input in as many places as possible.

Although we achieved a mathematically (and intuitively) correct result in this case, it is unfortunate that there appear to be many local minima into which the solution could fall. Some of the local minima could be close to an actual global minimum and could thereby yield a solution that is useful for the application at hand. In other cases a local-minimum solution, while very good mathematically, might be counterintuitive and thus unacceptable for a particular application. This unfortunate circumstance stems, in part, from the fact that there is no unique global minimum for the L^0-norm problem. This differs from the uniquely solvable L^2-norm problem. Convergence to a local minimum is guaranteed; convergence to a global minimum is not.

Figure 5 shows the unweighted least-squares solution (rewrapped) for comparison with Fig. 4(c). Note that the least-squares solution underestimates the magnitude of the phase gradients everywhere.

We now illustrate convergence to a local minimum in which the resulting solution is intuitively undesirable. Figure 6 shows a rotated and translated wrapped phase shear. Figures 7 and 8 show a sequence similar to that of Figs. 2–4. Because the input is nothing more than a rotated and translated version of the previous phase-shear

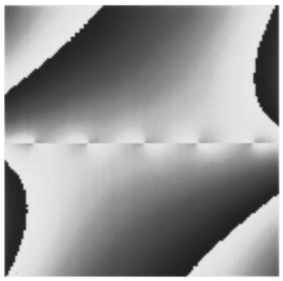

Fig. 5. Rewrapped unweighted least-squares solution to the phase-shear problem.

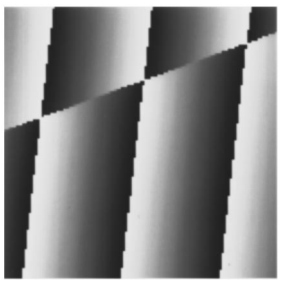

Fig. 6. Rotated and translated version of the wrapped phase-shear data.

problem, we would expect the solution to be a rotated and translated version of the previous solution, shown in Fig. 4(f).

Note, however, that we did not obtain our intuitively correct answer; namely, the rotated and translated version of Fig. 4(f). Instead, we obtained the result shown in

Fig. 8(f). Even though the partially converged results shown in Figs. 7(c), 7(f), and 8(c) looked as if they were progressing to the intuitively correct solution, the data-dependent weight progression did not allow this to happen. Instead, the partially converged solution and the data-dependent weights pinched off the continuous solution along a nearly vertical shear line as shown in Fig. 8(d). Thus the final converged result is simply the solution at a local minimum, not the minimum L^0-norm solu tion as desired. The result shown in Fig. 8(f), while not the global optimum, is still quite good in the mathematical sense. Specifically, this solution has gradients that-differ from the input in only 115 locations. We would have preferred the final weights to lie entirely along the

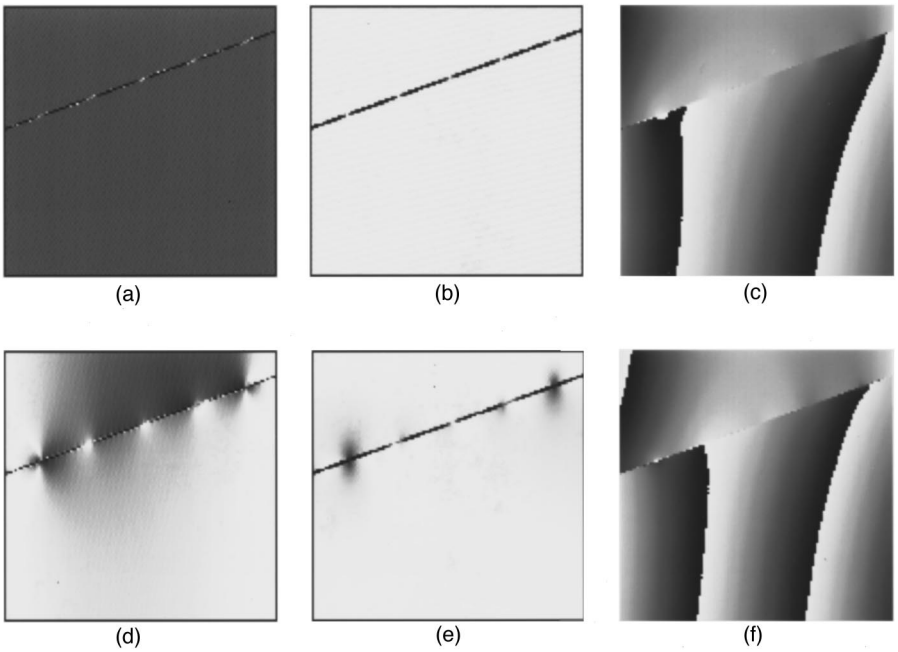

Fig. 7. Progress of Algorithm LP-NORM on the rotated and translated phase-shear problem: (a) data-dependent x-gradient weights, $U_0(i, j)$, from the initial guess, $\phi(i, j) = 0$; (b) data-dependent y-gradient weights, $V_0(i, j)$, from the initial guess, $\phi(i, j) = 0$; (c) rewrapped solution after iteration $l = 1$; (d) $U_1(i, j)$; (e) $V_1(i, j)$; (f) rewrapped solution after iteration $l = 2$.

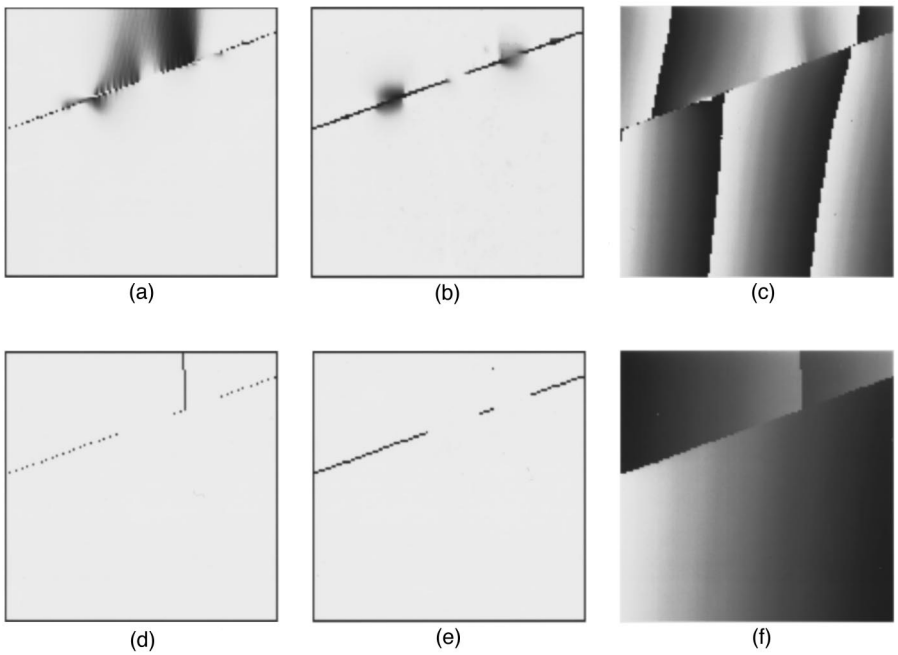

Fig. 8. Further progress and final convergence on the rotated phase-shear problem: (a) $U_4(i, j)$. (b) $V_4(i, j)$. (c) Rewrapped solution after iteration $l = 5$. (d) x-gradient weights, $U_{17}(i, j)$, after convergence to a local minimum. (e) y-gradient weights, $V_{17}(i, j)$, after convergence to a local minimum. (f) Unwrapped result after iteration $l = 17$. Remember that the rewrapped solution at this point agrees exactly with the original wrapped input.

rotated shear line and to provide the globally optimal and intuitively correct solution, but this, of course, did not happen.

Even if the algorithm were to converge to one of several identical (in norm) global minima, the solution might be undesirable. It might be undesirable because it could strongly violate our preconceived notion of correctness or could violate some other physical property required for the specific application.

One such example in which the true minimum L^0-norm solution might be undesirable is shown by unwrapping the shifted phase-shear data shown in Fig. 9. Again, we perceive the intuitively correct horizontal shear occurring along the line 13 samples below the upper border. The minimum L^0-norm solution and the data-dependent weights are shown at convergence ($l = 9$) in Fig. 10. It is interesting that this solution, while mathematically optimal, does not support our preconceived notion that the shear should be maintained solely along the perceived horizontal shear line. In fact, this is one of two possible mathematically optimal solutions, neither of which supports the horizontal shear hypothesis. [The second optimal solution results from the x-gradient weights containing a zero-weight line emanating from the leftmost residue to the top border. Thus the x-gradient weights would consist of four equal-length vertical line segments instead of the three segments shown in Fig. 10(a). The y-gradient weights would consist of only one rightmost horizontal line segment instead of the two segments shown in Fig. 10(b).] The 64 nonunity data-dependent weights begin at the residue locations and proceed in lines to the boundaries as shown in Figs. 10(a) and 10(b). In order to maintain the intuitively correct horizontal shear, more than 64 nonunity weights would be required. If the horizontal shear line were desired or required for the particular application at hand, this minimum L^0-norm solution would be undesirable. It should be emphasized that the raw wrapped data alone cannot support any preconceived notion of correctness. Because residues exist, many solutions are possible. Global or local minimum L^0-norm solutions sometimes give intuitively correct and usable results; at other times they do not.

We now present some examples drawn from actual applications. Figure 11(a) shows the wrapped phase from a MRI experiment.[40] The 256×256 sample wrapped phase must be unwrapped to allow correct water/fat separation.[13–15] Briefly, the MRI water/fat separation problem is described as follows. A MRI signal of the form

$$s = [W + F \exp(j\Phi)]\exp[j(\theta_0 + \varphi)] \tag{45}$$

is obtained by conventional procedures with certain parameters subject to operator control. Here W and F are water and fat signal intensities, respectively. Φ is a water/fat chemical phase shift, θ_0 is a constant phase for a given pixel, and φ is a phase shift that is due to inhomogeneities in the magnetic field. The operator can control Φ to obtain the following two signals:

$$s_0 = (W + F)\exp(j\theta_0), \tag{46}$$

$$s_1 = (W - F)\exp[j(\theta_0 + \varphi)]. \tag{47}$$

It is easily shown that

$$\frac{\bar{s}_0 s_1}{s_0 \bar{s}_1} = \exp(j2\varphi). \tag{48}$$

where the overbar represents complex conjugation.

We can extract wrapped values of phase, denoted as ψ, from Eq. (48). If ψ can be unwrapped, we can estimate 2φ. The unwrapped estimate is then divided by 2 to yield the estimate for φ, denoted as $\hat{\varphi}$. Finally, we compute

$$I_0 = |0.5[s_0 + s_1 \exp(-j\hat{\varphi})]|, \tag{49}$$

$$I_1 = |0.5[s_0 - s_1 \exp(-j\hat{\varphi})]|. \tag{50}$$

Fig. 9. Another phase surface with shear. The perceived horizontal shear line is within 13 pixels of the top border. This data set is simply a vertically shifted version of the data shown in Fig. 1(a). The corresponding residues [as in Fig. 1(b)] shift accordingly.

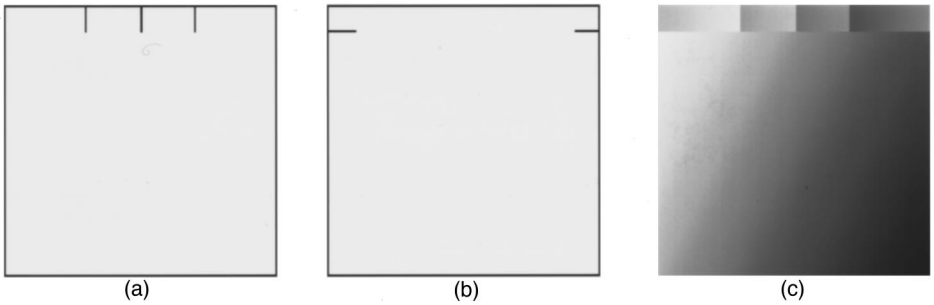

Fig. 10. Converged solution for the vertically shifted phase-shear problem: (a) $U_9(i, j)$. (b) $V_9(i, j)$. (c) Minimum L^0-norm solution after $l = 9$ iterations. This is one of two possible mathematically optimal solutions. Neither mathematically optimal solution maintains the phase shear solely along the horizontal shear line. There are only 64 locations in which the solution gradients disagree with the input data gradients. In order to maintain the intuitively correct horizontal shear, more than 64 low-valued weights would be required. This example shows that the mathematically optimal solution may not be what we desire in practice.

One of the two images, I_0 or I_1, will be the water image, and the other will be the fat image. Successful water/fat separation depends on successful phase unwrapping.

Because of the many residues in this example [see Fig. 11(b)], unwrapping is problematic for path-following and weighted least-squares algorithms because branch cuts must be generated from the residue map and/or weights must be externally supplied.

The unwrapped phase shown in Fig. 12(a) is an acceptable result because it allowed correct water/fat separation. It is not known whether the unwrapped phase represents an actual minimum L^0-norm solution or just a local-minimum solution. Nevertheless, this useful result was achieved without the explicit use of branch cuts or of externally supplied weights. Note that the data-dependent weights shown in Figs. 12(b) and 12(c) have low values in regions where residues exist, as expected [see Fig. 11(b)].

A MR example in which correct water/fat separation was not achieved is shown in Fig. 13. Although the unwrapping appeared qualitatively correct, there are isolated regions where the phase is shifted by $\pm 2\pi$ from the values required for correct water/fat separation. Apparently, this is a case in which the correct solution is just not achievable unambiguously from the input data. It appears to be a case of wanting φ but getting ϕ instead. Path-following unwrapping methods also failed to produce the correct unwrapped phase for this example.

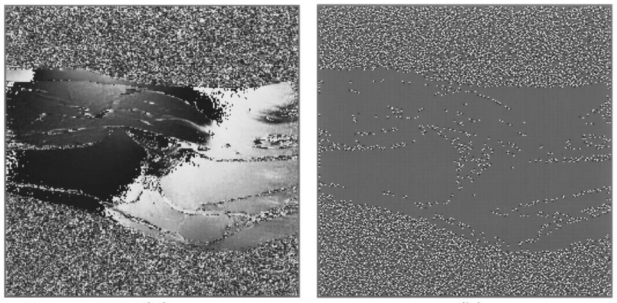

Fig. 11. Magnetic resonance example: (a) Wrapped phase of MR knee image for use in the water/fat separation problem. These data must be successfully unwrapped to allow correct water/fat separation. (b) Residue map containing 5620 positive and 5620 negative residues. Most of the residues come from the noise region outside the boundary of the knee.

The preceding example provides, however, one of the most visually interesting sequence of weight progressions of the experiments performed thus far. The data-dependent weights for the first four iterations are shown in Fig. 14. The weights at convergence ($l = 8$) are shown in Fig. 15.

One final example, drawn from SAR interferometry, is shown in Fig. 16. The 512×512 sample wrapped phase (representing a scaled and wrapped version of actual terrain elevation) is shown in Fig. 16(a). There are several regions visible in the wrapped phase where the fringes are highly irregular, broken up, and noisy. These are the regions where most of the 2937 positive and 2940 negative residues are concentrated. These regions of noisy phase result from low signal-to-noise ratios (e.g., shadows and low radar return) in the SAR image pairs used to construct the interferometric phase values shown. The unwrapped result (with a planar trend removed to allow easier visualization) is shown in Fig. 16(b). The algorithm converged in ten iterations and appeared to produce a useful and qualitatively correct unwrapped phase. Notice the unwrapped phase estimates in the noisy regions. The data-dependent weights effectively isolated these noisy regions from the remainder of the unwrapped data. The data-dependent weights at convergence ($l = 10$) are shown in Figs. 16(c) and 16(d). The rewrapped unweighted least-squares solution for the SAR interferometry example is shown in Fig. 17 and should be compared with that in Fig. 16(a). Remember that the correct solution for the SAR interferometry problem is obtained from Algorithm LP-NORM and, when rewrapped, is exactly the same as that shown in Fig. 16(a).

5. CONCLUSION

As a generalization of weighted least squares, we have presented an algorithm for minimum L^p-norm 2-D phase unwrapping. While computationally intense because of its doubly iterative structure, it can and does produce useful solutions for a variety of phase unwrapping problems. This new algorithm draws heavily on, and uses the methods devised for, transform-based unweighted and weighted least squares. In fact, Algorithm LP-NORM is dominated by the computation of forward and inverse transforms imbedded in the solution process of step 8, which is then called iteratively. The total number of for-

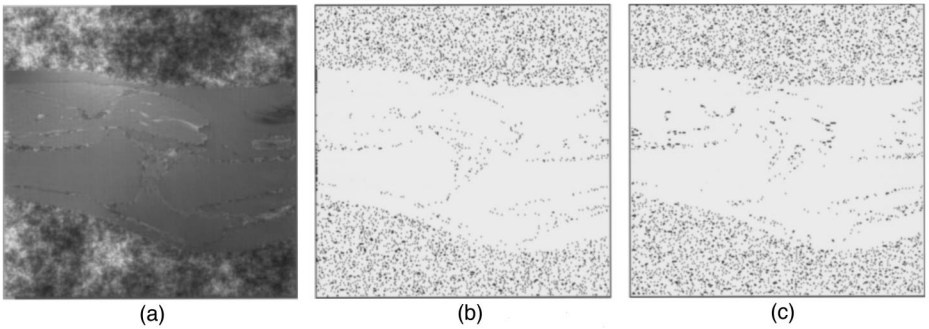

Fig. 12. Unwrapped result of MR knee data: (a) Unwrapped phase after convergence at $l = 11$ iterations. This unwrapped phase produced the correct water/fat separation. (b) x-gradient weights at convergence, $U_{11}(i, j)$. (c) y-gradient weights at convergence, $V_{11}(i, j)$.

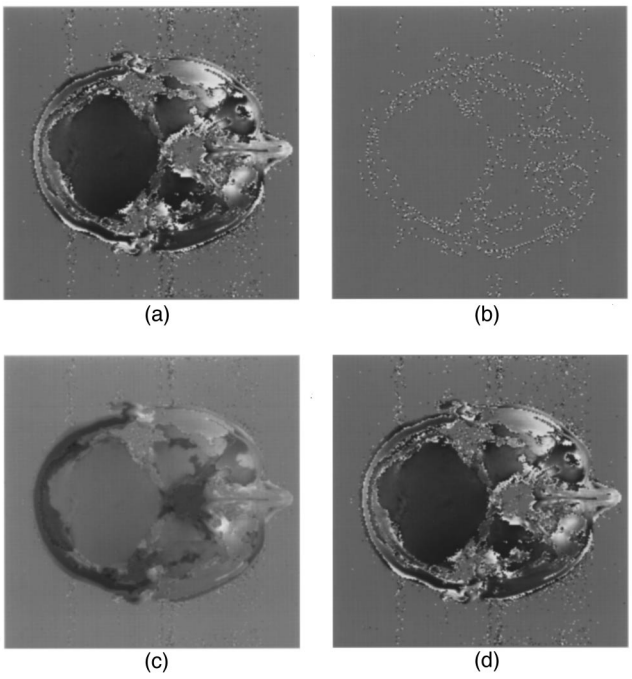

Fig. 13. Another MR example: (a) Wrapped phase of MR head image. (b) Residue map containing 963 positive and 963 negative residues. (c) Unwrapped phase at convergence ($l = 8$). (d) The rewrapped phase is identical to the input wrapped phase as expected. However, the unwrapped result did not produce the correct water/fat separation in all regions of the MR image.

ward and inverse transforms depends on the product of the maximum inner and outer iteration counts. Researchers who have successfully implemented the transform-based unweighted and weighted least-squares unwrapping algorithms will have no difficulty imbedding those methods in the new algorithm. While we believe that there is no best unwrapping technique, there could be a best algorithm for a particular application. It is in this spirit that our new algorithm was developed. We hope that other researchers implement it, employ it in their various applications, and improve on it. From continued research new ideas will spring forth in the quest for the "best" phase unwrapping method.

ACKNOWLEDGMENTS

This research was performed at Sandia National Laboratories, supported by the U.S. Department of Energy, under contract DE-AC04-94AL85000. We thank the anonymous reviewers for their thorough review and helpful suggestions.

APPENDIX A: REVIEW OF UNWEIGHTED AND WEIGHTED LEAST-SQUARES PHASE UNWRAPPING

Our motivations are to encourage other researchers to implement the algorithms that we present, to test them on their current applications, and to improve on them. Therefore we will review the unweighted and weighted least-squares methodology of Ref. 29 to establish a common notation with that developed in this paper as an aid to implementation. This will help the present paper to stand alone without requiring constant referral to previously published literature. Most of the material presented in the next two subsections is drawn from Ref. 29.

We will use both matrix notation for economy and, because the matrices are sparse, 2-D array notation to indicate the actual computations performed over indices (i, j). This should present no confusion because, as we will show below, all 2-D array quantities to be computed [e.g., $\phi(i, j)$, $U(i, j)$, and $V(i, j)$] are easily interpreted as one-dimensional (1-D) vectors when written in matrix

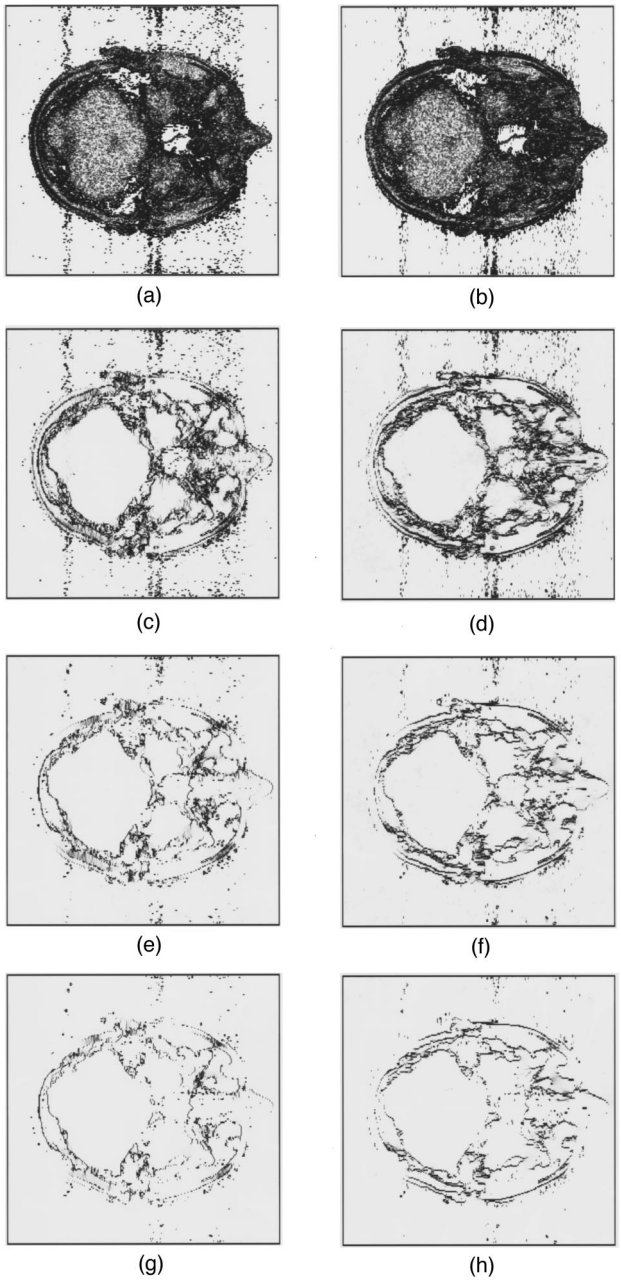

Fig. 14. Sequence of data-dependent weights for the first four iterations: (a) initial x-gradient weights ($l = 0$), (b) initial y-gradient weights ($l = 0$), (c) x-gradient weights ($l = 1$), (d) y-gradient weights ($l = 1$), (e) x-gradient weights ($l = 2$), (f) y-gradient weights ($l = 2$), (g) x-gradient weights ($l = 3$), (h) y-gradient weights ($l = 3$).

notation. (The 1-D vectors are obtained by concatenation of columns of the corresponding 2-D array.) All 2-D matrices are operators that operate on 1-D vectors. We will never write 2-D matrices explicitly in array notation. This notation should become clear in the following review.

1. Unweighted Least-Squares Two-Dimensional Phase Unwrapping

Unweighted least-squares phase unwrapping amounts to solving Poisson's equation [Eq. (34)] with Neumann boundary conditions [Eqs. (2) and (3)]. If we write the discrete version of Poisson's equation in matrix notation, we must solve

$$\mathbf{P}\phi = \rho, \tag{A1}$$

where $\mathbf{P}$ is a matrix operator that performs the discrete Laplacian operation on the vector ϕ. $\mathbf{P}\phi$, when written in 2-D array notation, yields the left-hand side of Eq. (34). Again, we emphasize that the 1-D vector, ρ, when written in 2-D array notation, is Eq. (35). We solve the unweighted least-squares phase unwrapping embodied in Poisson's equation by using 2-D discrete cosine transforms (DCT's)[41] as follows:

Algorithm LS (least-squares 2-D phase unwrapping):

1. Perform the 2-D forward DCT of the array of values $\rho_{i,j}$ to yield the 2-D DCT values $\hat{\rho}_{i,j}$.
2. Compute $\hat{\phi}_{i,j} = \hat{\rho}_{i,j}/2[\cos(\pi i/M) + \cos(\pi j/N) - 2]$.
3. Perform the 2-D inverse DCT of $\hat{\phi}_{i,j}$ to obtain the least-squares unwrapped phase values $\phi_{i,j}$.

It is important to note that we cannot evaluate $\hat{\phi}_{i,j}$ in step 2 above for both $i = 0$ and $j = 0$ simultaneously because the denominator becomes zero. $\hat{\phi}_{0,0}$ is indeterminate because Poisson's equation cannot be solved for a constant bias component (i.e., the matrix $\mathbf{P}$ is singular). In practice, we usually set $\hat{\phi}_{0,0} = \hat{\rho}_{0,0}$ to leave the bias unchanged.

Algorithm LS is also easily implemented with the use of fast Fourier transforms.[42]

2. Weighted Least-Squares Two-Dimensional Phase Unwrapping

Assuming that we have a matrix $\mathbf{W}$ that contains an explicit set of weights that apply to the x and y phase gradients (differences), we obtain the normal equations for the weighted least-squares problem as

$$\mathbf{A}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{A}\phi = \mathbf{c}. \tag{A2}$$

Now let $\mathbf{Q} = \mathbf{A}^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\mathbf{W}\mathbf{A}$, and substitute this into Eq. (A2) to obtain

$$\mathbf{Q}\phi = \mathbf{c}, \tag{A3}$$

where $\mathbf{c}$ is nothing more than a vector formed from the modified discrete Laplacian (to be defined below) of the weighted wrapped phase differences.

Here again, we consider $\mathbf{Q}$ as a matrix that performs the modified discrete Laplacian operation on the vector ϕ.

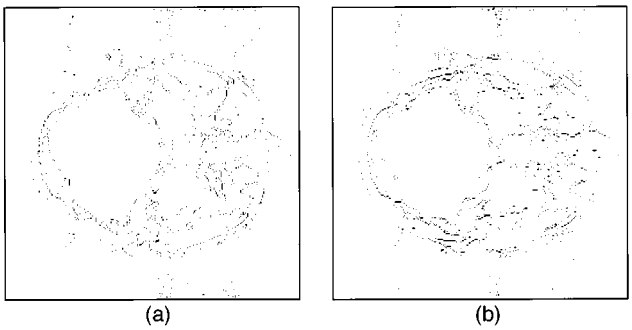

Fig. 15. Data-dependent weights at convergence ($l = 8$): (a) x-gradient weights, (b) y-gradient weights.

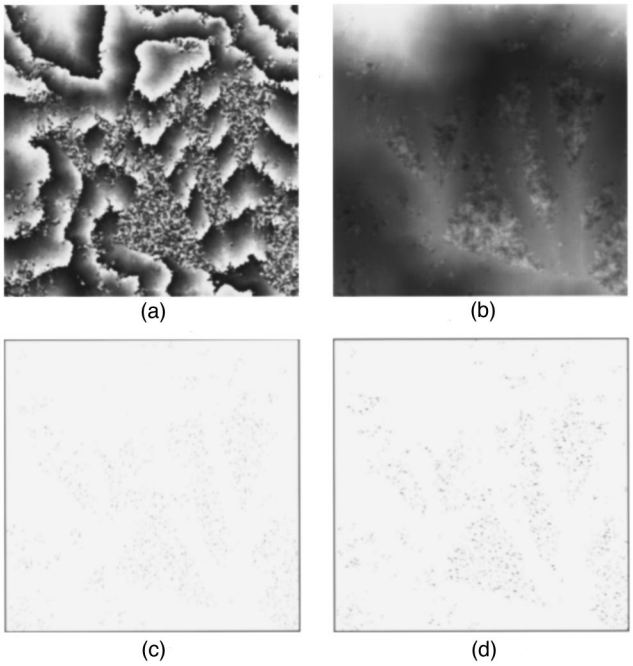

Fig. 16. A SAR interferometry example: (a) Wrapped phase (512×512) representing a scaled and wrapped version of actual terrain elevation. This image also represents the rewrapped solution of Algorithm LP-NORM at convergence. (b) Unwrapped result (with planar trend removed to allow easier visualization). The algorithm converged after $l = 10$ iterations. (c) x-gradient weights at convergence. (d) y-gradient weights at convergence.

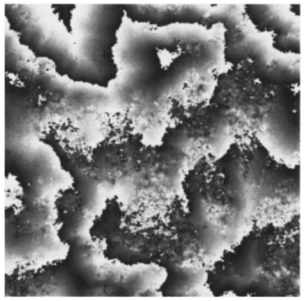

Fig. 17. (Rewrapped) unweighted least-squares solution for the SAR interferometry example. This least-squares solution has underestimated the magnitude of the phase gradients everywhere. The correct (rewrapped) solution for this SAR problem comes from Algorithm LP-NORM and is shown in Fig. 16(a).

Specifically, if we expand Eq. (A3) in array notation, we obtain

$$(\phi_{i+1,j} - \phi_{i,j})\tilde{U}(i,j) + (\phi_{i,j+1} - \phi_{i,j})\tilde{V}(i,j) - (\phi_{i,j} - \phi_{i-1,j})\tilde{U}(i-1,j) - (\phi_{i,j} - \phi_{i,j-1})\tilde{V}(i,j-1) = c(i,j), \quad \text{(A4)}$$

where

$$c(i,j) = f_{i,j}\tilde{U}(i,j) - f_{i-1,j}\tilde{U}(i-1,j) + g_{i,j}\tilde{V}(i,j) - g_{i,j-1}\tilde{V}(i,j-1) \quad \text{(A5)}$$

and $\tilde{U}(i,j)$ and $\tilde{V}(i,j)$ are arbitrarily defined weights applied to the measured wrapped phase differences $f_{i,j}$ and $g_{i,j}$, respectively. Note the equivalence in form between, on the one hand, Eqs. (A4) and (A5) and, on the other hand, Eq. (31), except that the weights $U(i,j)$ and $V(i,j)$ [Eqs. (32) and (33), respectively] are data dependent when $p \neq 2$. Thus we consider Eqs. (A4) and (A5) to represent the modified discrete Laplacian operation, whether the weights are data independent or data dependent.

In Ref. 29 we presented a preconditioned conjugate gradient algorithm for solving Eq. (A3). We repeat the algorithm below (now designated Algorithm WLS) for convenience.

Algorithm WLS (weighted least-squares 2-D phase unwrapping):

1. $k = 0$, $\phi_0 = \phi_I$, $\mathbf{r}_0 = \mathbf{c} - \mathbf{Q}\phi_I$: ϕ_I = initial guess; otherwise $\phi_I = 0$.
2. While ($\mathbf{r}_k \neq 0$) solve $\mathbf{P}\mathbf{z}_k = \mathbf{r}_k$ by using Algorithm LS.
3. $k = k + 1$.
4. If $k = 1$, $\mathbf{p}_1 = \mathbf{z}_0$.
5. If $k > 1$, then $\beta_k = \mathbf{r}_{k-1}^{\mathrm{T}}\mathbf{z}_{k-1}/\mathbf{r}_{k-2}^{\mathrm{T}}\mathbf{z}_{k-2}$ and $\mathbf{p}_k = \mathbf{z}_{k-1} + \beta_k\mathbf{p}_{k-1}$.
6. Perform one scalar and two vector updates: $\alpha_k = \mathbf{r}_{k-1}^{\mathrm{T}}\mathbf{z}_{k-1}/\mathbf{p}_k^{\mathrm{T}}\mathbf{Q}\mathbf{p}_k$, $\phi_k = \phi_{k-1} + \alpha_k\mathbf{p}_k$, and $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k\mathbf{Q}\mathbf{p}_k$.
7. If $k \geqslant k_{\max}$, or $\|\mathbf{r}_k\| < \epsilon\|\mathbf{r}_0\|$, end; otherwise go to step 2.

This is an iterative algorithm that allows us to incorporate the DCT method efficiently within the robust structure of conjugate gradients to solve the weighted least-squares problem. It should be noted that unweighted and weighted least-squares phase unwrapping also can be solved with multigrid methods.[37,38] It was noted in Ref. 30 that the matrix $\mathbf{Q}$ is singular, which can cause the algorithm to diverge. Possible divergence can be completely avoided, in practice, if any constant biases (average values) are removed (at each iteration k) from the vectors $\mathbf{r}$, $\mathbf{z}$, $\mathbf{p}$, and ϕ. This simple procedure prevents the arbitrary constant zero-eigenvalue solution from growing without bound and swamping the true solution. With this simple modification we have never experienced nonconvergent behavior of Algorithm WLS.

REFERENCES AND NOTES

1. L. C. Graham, "Synthetic interferometer radar for topographic mapping," Proc. IEEE **62**, 763–768 (1974).
2. R. M. Goldstein, H. A. Zebker, and C. L. Werner, "Satellite radar interferometry: two-dimensional phase unwrapping," Radio Sci. **23**, 713–720 (1988).
3. H. A. Zebker and R. M. Goldstein, "Topographic mapping from interferometric synthetic aperture radar observations," J. Geophys. Res. **91**, 4993–4999 (1986).
4. F. K. Li and R. M. Goldstein, "Studies of multibaseline spaceborne interferometric synthetic aperture radars," IEEE Trans. Geosci. Remote Sens. **28**, 88–97 (1990).
5. A. L. Gray and P. J. Farris-Manning, "Repeat-pass interferometry with airborne synthetic aperture radar," IEEE Trans. Geosci. Remote Sens. **31**, 180–191 (1993).
6. P. H. Eichel, D. C. Ghiglia, C. V. Jakowatz, Jr., P. A. Thompson, and D. E. Wahl, "Spotlight SAR interferometry for terrain elevation mapping and interferometric change detection," Rep. SAND93-2072J (Sandia National Laboratories, Albuquerque, N.M., 1993).
7. D. Massonnet, M. Rossi, C. Carmona, F. Adragna, G. Peltzer, K. Feigl, and T. Rabaute, "The displacement field of the Landers earthquake mapped by radar interferometry," Nature (London) **364**, 138–142 (1993).
8. K. J. Gåsvik, *Optical Metrology*, 2nd ed. (Wiley, Chichester, UK, 1995).
9. D. W. Robinson and G. T. Reid, *Interferogram Analysis: Digital Fringe Pattern Measurement Techniques* (Institute of Physics, Bristol, UK, 1993).
10. D. L. Fried, "Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements," J. Opt. Soc. Am. **67**, 370–375 (1977).
11. R. H. Hudgin, "Wave-front reconstruction for compensated imaging," J. Opt. Soc. Am. **67**, 375–378 (1977).
12. R. J. Noll, "Phase estimates from slope-type wave-front sensors," J. Opt. Soc. Am. **68**, 139–140 (1978).
13. G. H. Glover and E. Schneider, "Three-point Dixon technique for true water/fat decomposition with B_0 inhomogeneity correction," Magn. Reson. Med. **18**, 371–383 (1991).
14. J. Szumowski, W. R. Coshow, F. Li, and S. F. Quinn, "Phase unwrapping in the three-point Dixon method for fat suppression MR imaging," Radiology, **192**, 555–561 (1994).
15. 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. **4**, 667–676 (1995).
16. K. Itoh, "Analysis of the phase unwrapping problem," Appl. Opt. **21**, 2470 (1982).
17. D. C. Ghiglia, G. A. Mastin, and L. A. Romero, "Cellular automata method for phase unwrapping," J. Opt. Soc. Am. A **4**, 267–280 (1987).
18. J. M. Huntley, "Noise-immune phase unwrapping algorithm," Appl. Opt. **28**, 3268–3270 (1989).
19. D. J. Bone, "Fourier fringe analysis: the two-dimensional phase unwrapping problem," Appl. Opt. **30**, 3627–3632 (1991).
20. R. Cusack, J. M. Huntley, and H. T. Goldrein, "Improved noise-immune phase-unwrapping algorithm," Appl. Opt. **34**, 781–789 (1995).
21. J. A. Quiroga, A. González-Cano, and E. Bernabeu, "Stable-marriages algorithm for preprocessing phase maps with discontinuity sources," Appl. Opt. **34**, 5029–5038 (1995).
22. J. R. Buckland, J. M. Huntley, and S. R. E. Turner, "Unwrapping noisy phase maps by use of a minimum-cost-matching algorithm," Appl. Opt. **34**, 5100–5108 (1995).
23. N. H. Ching, D. Rosenfeld, and M. Braun, "Two-dimensional phase unwrapping using a minimum spanning tree algorithm," IEEE Trans. Image Process. **1**, 355–365 (1992).
24. B. R. Hunt, "Matrix formulation of the reconstruction of phase values from phase differences," J. Opt. Soc. Am. **69**, 393–399 (1979).
25. H. Takajo and T. Takahashi, "Least-squares phase estimation from phase differences," J. Opt. Soc. Am. A **5**, 416–425 (1988).
26. H. Takajo and T. Takahashi, "Noniterative method for obtaining the exact solution for the normal equation in least-squares phase estimation from the phase difference," J. Opt. Soc. Am. A **5**, 1818–1827 (1988).
27. D. C. Ghiglia and L. A. Romero, "Direct phase estimation

from phase differences using fast elliptic partial differential equation solvers," Opt. Lett. **14**, 1107–1109 (1989).
28. B. L. Busbee, G. H. Gollub, and C. W. Nielson, "On direct methods for solving Poisson's equations," SIAM J. Numer. Anal. **7**, 627–656 (1970).
29. 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. Am. A **11**, 107–117 (1994).
30. J. L. Marroquin and M. Rivera, "Quadratic regularization functionals for phase unwrapping," J. Opt. Soc. Am. A **12**, 2393–2400 (1995).
31. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes: The Art of Scientific Computing*, 2nd ed. (Cambridge U. Press, Cambridge, 1992), Chap. 18.
32. D. Just, Deutsche Forschungsanstalt für Luft-und Raumfahrt (DLR), Wessling, Germany, July 1995 (personal communication).
33. D. Just, N. Adam, M. Schwäbisch, and R. Bamler, "Comparison of phase unwrapping algorithms for SAR interferograms," presented at the International Geoscience and Remote Sensing Symposium (IGARSS '95), Firenze, Italy, July 10–14, 1995.
34. M. D. Pritt, Loral Federal Systems, Gaithersburg, Md., September 1995 (personal communication).
35. G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 4th ed. (Academic, San Diego, 1995), Chap. 17.
36. G. M. Ewing, *Calculus of Variations with Applications* (Dover, New York, 1985).
37. M. D. Pritt, "Phase unwrapping by means of multigrid techniques for inteferometric SAR," IEEE Trans. Geosci. Remote Sens. **34**, 728–738 (1996).
38. M. D. Pritt, "Multigrid phase unwrapping for interferometric SAR," in Proceedings of the IEEE International Geoscience Remote Sensing Symposium (IEEE, **Au:city**, 1995), Vol. 1, pp. 562–564.
39. J. M. Huntley and J. R. Buckland, "Characterization of sources of 2π phase discontinuity in speckle interferograms," J. Opt. Soc. Am. A **12**, 1990–1996 (1995).
40. The wrapped phase data sets shown in Fig. 11 and below in Fig. 13 are courtesy of Gang Zhu. They come from the ESTEEM clinical MRI scanner manufactured by Elscint MR Inc., Fort Collins. Colo.
41. J. S. Lim, "The discrete cosine transform," in *Two-Dimensional Signal and Image Processing* (Prentice-Hall, Englewood Cliffs, N.J., 1990), pp. 148–157.
42. M. D. Pritt and J. S. Shipman, "Least-squares two-dimensional phase unwrapping using FFTs," IEEE Trans. Geosci. Remote Sens. **32**, 706–708 (1994).