Phase Unwrapping via Graph Cuts

José M. Bioucas-Dias, *Senior Member, IEEE*, and Gonçalo Valadão

***Abstract*—Phase unwrapping is the inference of absolute phase from modulo-2π phase. This paper introduces a new energy minimization framework for phase unwrapping. The considered objective functions are first-order Markov random fields. We provide an exact energy minimization algorithm, whenever the corresponding clique potentials are convex, namely for the phase unwrapping classical L^p norm, with $p \geq 1$. Its complexity is $KT(n, 3n)$, where K is the length of the absolute phase domain measured in 2π units and $T(n, m)$ is the complexity of a max-flow computation in a graph with n nodes and m edges. For nonconvex clique potentials, often used owing to their discontinuity preserving ability, we face an NP-hard problem for which we devise an approximate solution. Both algorithms solve integer optimization problems by computing a sequence of binary optimizations, each one solved by graph cut techniques. Accordingly, we name the two algorithms PUMA, for phase unwrapping max-flow/min-cut. A set of experimental results illustrates the effectiveness of the proposed approach and its competitiveness in comparison with state-of-the-art phase unwrapping algorithms.**

***Index Terms*—Computed image, discontinuity preservability, energy minimization, graph cuts, image reconstruction, InSAR, integer optimization, magnetic resonance imaging (MRI), phase unwrapping (PU), submodularity.**

I. Introduction

THE need for phase estimation is common to many imaging techniques, from which we point up interferometric synthetic aperture radar and sonar (InSAR/InSAS) [3]–[9], magnetic resonance imaging (MRI) [10], [11], and optical interferometry [12]. In InSAR, as in InSAS, two or more antennas measure the phase between them and the terrain; the topography may then be inferred from the difference between those phases, relying on simple geometric reasoning. In MRI phase is used, namely, to determine magnetic field deviation maps, which are used to correct echo-planar image geometric distortions [13], to determine chemical shift-based thermometry [14], and to implement BOLDcontrast-based venography [15]. In optical interferometry, phase measurements are used to detect objects shape, deformation, and vibration [12].

In all the examples above, the acquisition system can only measure phase modulo-2π, the so-called *principal* phase value, or *wrapped* phase. Formally, we have

$$\phi = \psi + 2k\pi \tag{1}$$

Manuscript received October 21, 2005; revised August 9, 2006. This work was supported in part by the Fundação para a Ciência e Tecnologia under the project PDCTE/CPS/49967/2003, in part by the European Space Agency under the Project ESA/C1:2422/2003, and in part by the Instituto de Telecomunicações under Project IT/LA/325/2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Robert P. Loce.

The authors are with the Instituto de Telecomunicações, Instituto Superior Técnico, and the Technical University of Lisbon, 1049-001 Lisboa, Portugal (e-mail: bioucas@lx.it.pt; gvaladao@lx.it.pt).

Digital Object Identifier 10.1109/TIP.2006.888351

where ϕ is the true phase value (the so-called *absolute phase* value), ψ is the measured (wrapped) *modulo*-2π phase value, and $k \in \mathbb{Z}$ ($\mathbb{Z}$ denotes the set of integers) is an integer accounting for the number of 2π multiples [5].

Phase unwrapping (PU) is the process of recovering the absolute phase ϕ from the wrapped phase ψ. This is, however, an ill-posed problem, if no further information is added. In fact, an assumption taken by most phase unwrapping algorithms is that the absolute value of phase differences between neighboring pixels is less than π, the so-called *Itoh* condition [16]. If this assumption is not violated, the absolute phase can be easily determined, up to a constant. Itoh condition might be violated if the true phase surface is discontinuous, or if the wrapped phase is noisy. In either cases, PU becomes a very difficult problem, to which much attention has been devoted [5], [7], [17]–[20], [21].

Phase unwrapping approaches belong mainly to one of the following classes: path following [17], [22], [23], minimum L^p norm [19], [20], [24]–[26], [7], Bayesian/regularization [7], [18], [27]–[32], and parametric modeling [33], [34].

Path following algorithms apply line integration schemes over the wrapped phase image, and basically rely on the assumption that Itoh condition holds along the integration path. Wherever this condition fails, different integration paths may lead to different unwrapped phase values. Techniques employed to handle these inconsistencies include the so-called *branch cuts* [17] and *quality maps* [5, Ch. 4].

Minimum norm methods try to find a phase solution ϕ for which the L^p norm of the difference between absolute phase differences and wrapped phase differences (so a second-order difference) is minimized. This is, therefore, a global minimization in the sense that all the observed phases are used to compute a solution. With $p = 2$, we have a least squares method [35]. The exact solution with $p = 2$ is developed in [7] using network programming techniques. An approximation to the least squares solution can be obtained by relaxing the discrete domain $\mathbb{Z}^{\mathrm{MN}}$ to $\mathbb{R}^{\mathrm{MN}}$, where M and N are, respectively, the number of lines and columns, and applying FFT- or DCT-based techniques [5, Ch. 5], [24]. A drawback of the L^2 norm-based criterion is that it tends to smooth discontinuities, unless they are provided as binary weights. L^1 norm performs better than L^2 norm in what discontinuity preserving is concerned. Such a criterion has been solved exactly by Flynn [19] and Costantini [20], using network programming concepts. With $0 \leq p < 1$ the discontinuity preserving ability is further increased at stake, however, of highly complex algorithms [29], [31]. In particular, L^0 norm is generally accepted as the most desirable in practice. The minimization of L^0 norm is, however, an NP-hard problem [29], for which approximate algorithms have been proposed in [5, Ch. 5] and [26].

The Bayesian approach relies on a data-observation mechanism model, as well as *a priori* knowledge of the phase to be modeled. For instance, in [36], a nonlinear optimal filtering is

applied, while in [27], an InSAR observation model is considered, taking into account not only the image phase, but also the *backscattering coefficient* and *correlation factor* images, which are jointly recovered from InSAR image pairs; [37] proposes a fractal-based prior, and [32] employs dynamic programming techniques.

Finally, parametric algorithms constrain the unwrapped phase to a parametric surface. Low-order polynomial surfaces are used in [33]. Very often in real applications, just one polynomial is not enough to describe accurately the complete surface. In such cases the image is partitioned and different parametric models are applied to each partition [33].

A. Contributions

The main contribution of the paper is an energy minimization framework for phase unwrapping, where the minimization is carried out by a sequence of max-flow/min-cut calculations. The objective functions considered are first-order Markov random fields, with pairwise interactions. The associated energy is, therefore, a generalization of the classical L^p norm, used in phase unwrapping [25]. We show that, if the clique potentials are convex, the exact energy minimization is achieved by a finite sequence of binary minimizations, each one solved efficiently from the computational point of view, by a max-flow/min-cut calculation on a given graph, building on energy minimization results presented in [38], [39], and [40]; we, thus, benefit from existing efficient algorithms for graph max-flow/min-cut calculations [41]. Accordingly, we call the method to be presented PUMA algorithm (for PU-max-flow). Besides solving exactly the classical minimum L^p norm problem for $p \geq 1$, PUMA is able to minimize a wider class of energies, rendering flexibility to the method.

In image reconstruction and in phase unwrapping, in particular, it is well known that unknown discontinuities pose a challenging problem (as well as an usual one in practice), for which nonconvex clique potentials are critical to deal with [5, Ch. 5], [42]–[46]. Nonconvexity, however, turns our minimization problem into an NP-hard one [29], [39], and part of the concepts and results developed under the convexity assumption do not apply any more. Namely, energy cannot be minimized by a sequence of binary minimizations, nor each one of these problems can be solved by the former max-flow/min-cut calculations.

We also introduce an approximate algorithm that tackles those issues by 1) enlarging the configuration space of each binary problem and by 2) applying majorize minimize (MM) [47] concepts to our energy function, which still allow max-flow/min-cut calculations. For the sake of uniformity, we still term the obtained algorithm PUMA. Experimental results illustrate the state-of-the-art competitiveness of the presented algorithms.

After this paper was submitted, Darbon [48] and Kolmogorov [49], exploiting the concept of submodularity, independently generalized the class of energies herein studied, in the convex scenario. Besides pairwise terms depending on differences, they have included unary convex terms. These class of energies arise in many computer vision and image processing problems. The algorithms they propose are similar to ours, replacing a binary optimization by a sequence of two binary optimizations. A major contribution of Darbon and Kolmogorov is a tight bound on the number of steps.

Other related works are the steepest descent algorithm of Murota for minimizing $L^\#$ functions [50], the very fast algorithms of Darbon and Sigelle [51], Chambolle [52], and Hochbaum [53], if pairwise terms are absolute differences.

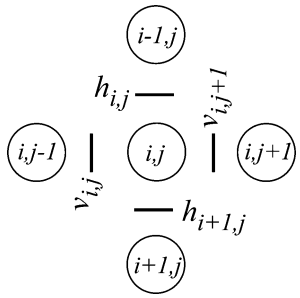

Fig. 1. Representation of the site (i, j) and its first-order neighbors along with the variables h_{ij} and v_{ij} signalling horizontal and vertical discontinuities, respectively.

II. Problem Formulation

Fig. 1 shows a site $(i, j) \in \mathbb{G}_0 \equiv \{(k, l) : k = 1, \ldots, M, l = 1, \ldots, N\}$ ($\mathbb{G}_0$ is the usual image pixel indexing 2-D grid) and its first-order neighbors along with the variables h_{ij} and v_{ij} signalling horizontal and vertical discontinuities, respectively; i.e., $h_{ij}, v_{ij} \in \{0, 1\}$ and $h_{ij}, v_{ij} = 0$ signals a discontinuity.

Let us define the energy

$$E(\mathbf{k}|\boldsymbol{\psi}) \equiv \sum_{ij \in \mathbb{G}_0} V\left(\Delta\phi_{ij}^h\right) v_{ij} + V\left(\Delta\phi_{ij}^v\right) h_{ij} \tag{2}$$

where $\mathbf{k} \equiv \{k_{ij} \in \mathbb{Z} : (i, j) \in \mathbb{G}_0\}$ is an image of integers, denoting 2π multiples, the so-called *wrap-count image*, $\boldsymbol{\psi} \equiv \{\psi_{ij} \in [-\pi, \pi) : (i, j) \in \mathbb{G}_0\}$ is the observed wrapped phase image, $V(\cdot)$ is the *clique potential*, a real-valued function,[1] and $(\cdot)^h$ and $(\cdot)^v$ denote pixel horizontal and vertical differences given by

$$\Delta\phi_{ij}^h \equiv \left[2\pi(k_{ij} - k_{ij-1}) - \Delta\psi_{ij}^h\right] \tag{3}$$

$$\Delta\phi_{ij}^v \equiv \left[2\pi(k_{ij} - k_{i-1j}) - \Delta\psi_{ij}^v\right] \tag{4}$$

$$\Delta\psi_{ij}^h \equiv \psi_{ij-1} - \psi_{ij} \tag{5}$$

$$\Delta\psi_{ij}^v \equiv \psi_{i-1j} - \psi_{ij}. \tag{6}$$

Furthermore, each of the above defined left-hand side terms (3)–(6) is defined to be zero if any of the right-hand side indexes [$(i, j-1)$ or $(i-1, j)$] does not belong to $\mathbb{G}_0$.

Our goal is to find the integer image $\mathbf{k}$ that minimizes energy (2), $\mathbf{k}$ being such that $\boldsymbol{\phi} = 2\pi\mathbf{k} + \boldsymbol{\psi}$, where $\boldsymbol{\phi}$ is the estimated unwrapped phase image. As will be seen in the next section, this energy minimization approach yields the classical minimum L^p norm formulation or a more general one, depending on the clique potential V.

We should stress that the variables h_{ij} and v_{ij}, conveying discontinuity information, are introduced when available. In PU jargon these images are the so-called *quality* maps. These maps

[1]The clique is a set of sites that are mutually neighbors. A clique function is a function defined on cliques, i.e., it depends only on site variables indexed by the respective clique elements.

can also be used as continuous variables in $[0,1]$, expressing prior knowledge on phase variability. Quality maps can be derived, for example, from correlation maps in InSAR, or from phase derivative variance in a more general setting [5, Ch.3]. Nevertheless, in practice, quality maps are often very noisy or unavailable, implying blind handling of discontinuities and, therefore, calling for nonconvex potentials.

Although energy (2) was introduced deterministically in a natural way, it can also be derived under a Bayesian perspective as in [7]. Consider a first-order Markov random field prior on the absolute phase image given by $p(\phi) = (1/Z)e^{-U(\phi)}$, where $U(\phi) \equiv \sum_{ij\in\mathbb{G}_0} V(\Delta\phi^h_{ij})v_{ij} + V(\Delta\phi^v_{ij})h_{ij}$, and Z is a normalizing constant. Assuming that the wrapped phase is noiseless, then $\phi = \psi + 2k\pi$. Therefore, the maximum *a posteriori* (MAP) estimate exactly amounts to minimize $E(\mathbf{k}|\boldsymbol{\psi})$ with respect to $\mathbf{k}$.

III. Energy Minimization by a Sequence of Binary Optimizations: Convex Potentials

In this section, we present in detail the PUMA algorithm. We show that for convex potentials V, the minimization of $E(\mathbf{k}|\boldsymbol{\psi})$ can be achieved through a sequence of binary optimizations; each binary problem is mapped onto a certain graph and a binary minimization obtained by computing a max-flow/min-cut on it. Finally, we address a set of potentials tailored to phase unwrapping.

A. Equivalence Between Local and Global Minimization

The following theorem is an extension of Lemma 1 in [7], which, in turn, is inspired by Lemma 1 of [19]. Assuming a convex clique potential V, it assures that if the minimum of $E(\mathbf{k}|\boldsymbol{\psi})$ is not yet reached, then, there exists a binary image $\boldsymbol{\delta} \in \{0,1\}^{\mathrm{MN}} \equiv \mathcal{B}$ (i.e., the elements of $\boldsymbol{\delta}$ are 0 or 1) such that $E(\mathbf{k}+\boldsymbol{\delta}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi})$. Therefore, if a given image $\mathbf{k}$ is locally optimal with respect to the neighborhood $\mathcal{N}_1(\mathbf{k}) \equiv \{\mathbf{k}+\boldsymbol{\delta} : \boldsymbol{\delta} \in \mathcal{B}\}$, i.e., if $E(\mathbf{k}'|\boldsymbol{\psi}) \geq E(\mathbf{k}|\boldsymbol{\psi})$ for all $\mathbf{k}' \in \mathcal{N}_1(\mathbf{k})$, then $\mathbf{k}$ it is also globally optimal.

Theorem 1: Let $\mathbf{k}_1$ and $\mathbf{k}_2$ be two wrap-count images such that

$$E(\mathbf{k_2}|\boldsymbol{\psi}) < E(\mathbf{k_1}|\boldsymbol{\psi}). \tag{7}$$

Then, if V is convex, there exists a binary image $\boldsymbol{\delta} \in \mathcal{B}$ such that

$$E(\mathbf{k_1}+\boldsymbol{\delta}|\boldsymbol{\psi}) < E(\mathbf{k}_1|\boldsymbol{\psi}). \tag{8}$$

Proof: See the Appendix.

B. Convergence Analysis

In accordance with Theorem 1, we can iteratively compute $\mathbf{k}^{t+1} = \mathbf{k}^t + \boldsymbol{\delta}$, where $\boldsymbol{\delta} \in \mathcal{B}$ is such that it minimizes[2] $E(\mathbf{k}^t+\boldsymbol{\delta}|\boldsymbol{\psi})$, until the the minimum energy is reached. There is of course the pertinent question of whether the algorithm stops and, if it does, in how many iterations. Assuming that $\boldsymbol{k}^0 = \mathbf{0}$, the next lemma, which is inspired in the Proposition 3.7 of [48], leads to the conclusion that after t iterations the algorithm minimizes $E(\cdot|\boldsymbol{\psi})$ in $\mathcal{D}_t \equiv \{\mathbf{k}' : 0 \leq k'_{ij} \leq t\}$.

Lemma 1: Let $\mathbf{k}^t$ be a globally optimal minimizer of $E(\cdot|\boldsymbol{\psi})$ on $\mathcal{D}_t$. Then, there exists an image $\mathbf{k}^{t+1}$ that is a global minimizer of $E(\cdot|\boldsymbol{\psi})$ on $\mathcal{D}_{t+1}$ and

$$\mathbf{k}^{t+1} - \mathbf{k}^t \in \mathcal{B}.$$

Therefore, $\mathbf{k}^{t+1}$ can be found by minimizing $E(\mathbf{k}+\boldsymbol{\delta}|\boldsymbol{\psi})$ with respect to $\boldsymbol{\delta} \in \mathcal{B}$.

Proof: See the Appendix.

Assume that the range of E spans over K wrap counts. Then its global minimizer is in the set $\mathcal{D}_{K-1}$, and, therefore, Lemma 1 assures that the iterative scheme

$$\begin{aligned} &\text{do} \quad \mathbf{k}^{t+1} = \arg\min_{\boldsymbol{\delta}\in\mathcal{B}} E(\mathbf{k}^t+\boldsymbol{\delta}|\boldsymbol{\psi}) \\ &\text{while } E(\mathbf{k}^{t+1}|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi}) \end{aligned}$$

starting with $\boldsymbol{k}^0 = \mathbf{0}$, finds this minimizer in at most K iterations. Its complexity is, therefore, KT, where T is the complexity of a binary optimization.

C. Mapping Binary Optimizations Onto Graph Max-Flows

Let $k^{t+1}_{ij} = k^t_{ij} + \delta_{ij}$ be the wrap count at time $t+1$ and pixel (i,j). Introducing k^{t+1}_{ij} into (3) and (4), we obtain, respectively

$$\Delta\phi^h_{ij} = \left[2\pi\left(k^{t+1}_{ij} - k^{t+1}_{ij-1}\right) - \Delta\psi^h_{ij}\right] \tag{9}$$

$$\Delta\phi^v_{ij} = \left[2\pi\left(k^{t+1}_{ij} - k^{t+1}_{i-1j}\right) - \Delta\psi^v_{ij}\right]. \tag{10}$$

After some simple manipulation, we get

$$\Delta\phi^h_{ij} = [2\pi(\delta_{ij} - \delta_{ij-1}) + a^h] \tag{11}$$

$$\Delta\phi^v_{ij} = [2\pi(\delta_{ij} - \delta_{i-1j}) + a^v] \tag{12}$$

where

$$a^h \equiv 2\pi\left(k^t_{ij} - k^t_{ij-1}\right) - \Delta\psi^h_{ij} \tag{13}$$

$$a^v \equiv 2\pi\left(k^t_{ij} - k^t_{i-1j}\right) - \Delta\psi^v_{ij}. \tag{14}$$

Now, introducing (11) and (12) into (2), we can rewrite energy $E(\mathbf{k}^t+\boldsymbol{\delta}|\boldsymbol{\psi})$ as a function of the binary variables $\delta_{ij} \in \{0,1\}$, i.e.,

$$\begin{aligned} E(\mathbf{k}^t+\boldsymbol{\delta}|\boldsymbol{\psi}) = \sum_{ij\in\mathbb{G}_0} &\underbrace{V\left[2\pi(\delta_{ij} - \delta_{ij-1}) + a^h\right] v_{ij}}_{E^{ij}_h(\delta_{ij-1},\delta_{ij})} \\ &+ \underbrace{V\left[2\pi(\delta_{ij} - \delta_{i-1j}) + a^v\right] h_{ij}}_{E^{ij}_v(\delta_{i-1j},\delta_{ij})}. \end{aligned} \tag{15}$$

Occasionally, and for the sake of notational simplicity, we use the representation

$$E(\mathbf{k}^t+\boldsymbol{\delta}|\boldsymbol{\psi}) = \sum_{ij\in\mathbb{G}_0} E^{ij}(\delta_i,\delta_j) \tag{16}$$

where indices i,j correspond now to the lexicographic column ordering of $\mathbb{G}_0$, $\delta_i \in \{0,1\}$, and $\boldsymbol{\delta} = \{\delta_i\} \in \{0,1\}^{\mathrm{MN}}$. Notice that with this representation some terms E^{ij} stand for horizontal cliques whereas others stand for vertical ones (e.g.,

[2]Or at least decreases.

E^{12} and $E^{1(M+1)}$ represent vertical and horizontal cliques, respectively).

The minimization of (15) with respect to $\boldsymbol{\delta}$ is now mapped onto a max-flow problem. Since the seminal work of Greig *et al.* [54], a considerable amount of research effort has been devoted to energy minimization via graph methods (see, e.g., [38]–[40] and [55]–[57]). Namely, the mapping of a minimization problem into a sequence of binary minimizations, computed by graph cut techniques, has been addressed in works [39] and [40]. Nevertheless, these two works assume the potentials to be either a metric or a semi-metric, which is not the case for the clique potentials that we are considering: from (15), it can be seen that $E^{ij} \neq E^{ji}$ as a consequence of the presence of a^h and a^v terms (by definition both a metric and a semi-metric satisfy the symmetry property). For this reason, we adopt the method proposed in [38], which generalizes the class of binary minimizations that can be solved by graph cuts. Furthermore, the graph structures therein proposed are simpler.

At this point a reference to work [57] should be made: it introduces an energy minimization for convex potentials also by computing a max-flow/min-cut on a certain graph. However, for a general convex potential that graph can be huge, imposing in practice, heavy computational and storage demands.

Following, then, [38], we now exploit a one-to-one map existing between the energy function (15) and cuts on a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ ($\mathcal{V}$ and $\mathcal{E}$ denote the set of vertices and edges, respectively) with nonnegative weights. The graph has two special vertices, namely the source s and the sink t. An $s-t$ cut $C = S, T$ is a partition of vertices $\mathcal{V}$ into two disjoint sets S and T, such that $s \in S$ and $t \in T$. The number of vertices is $2 + M \times N$ (two terminals, the source and the sink, plus the number of pixels). The cost of the cut is the sum of costs of all edges between S and T.

Using the notation above introduced, we have

$$\begin{aligned} E^{ij}(0,0) &= V(a)d_{ij} \\ E^{ij}(1,1) &= V(a)d_{ij} \\ E^{ij}(0,1) &= V(-2\pi + a)d_{ij} \\ E^{ij}(1,0) &= V(2\pi + a)d_{ij} \end{aligned} \tag{17}$$

where a represents a_h or a_v and d_{ij} represents h_{ij} or v_{ij}. Energy $E(\mathbf{k}^t + \boldsymbol{\delta}|\boldsymbol{\psi})$ is a particular case of the $\mathcal{F}^2$ class of functions addressed in [38], with zero unary terms. Roughly speaking,[3] a function of $\mathcal{F}^2$ is *graph representable*, i.e., there exists a one-to-one relation between configurations $\boldsymbol{\delta} \in \{0,1\}^{\mathrm{MN}}$ [i.e., points in the domain of $E(\mathbf{k}^t + \boldsymbol{\delta}|\boldsymbol{\psi})$] and $s-t$ cuts on that graph, if and only if holds

$$E^{ij}(0,0) + E^{ij}(1,1) \leq E^{ij}(0,1) + E^{ij}(1,0). \tag{18}$$

In terms of E^{ij} [see expression (17)] in equation (18) can be stated as $[V(-2\pi + a) + V(2\pi + a)]d_{ij} \geq 2V(a)d_{ij}$, which is verified due to convexity of V. So, our binary function is graph-representable.

[3]As defined in [38], a function E of n binary variables is called *graph-representable* if there exists a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with terminals s and t and a subset of vertices $\mathcal{V}_0 = \{v_1, \ldots, v_n\} \subset \mathcal{V} - \{s,t\}$ such that, for any configuration $\delta_1, \ldots, \delta_n$, the value of the energy $E(\delta_1, \ldots, \delta_n)$ is equal to a constant plus the cost of the minimum s-t-cut among all cuts $C = S, T$ in which $v_i \in S$, if $\delta_i = 0$, and $v_i \in T$, if $\delta_i = 1$ $(1 \leq i \leq n)$.

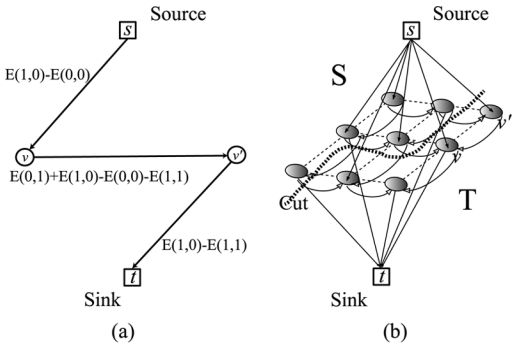

Fig. 2. (a) Elementary graph for a single energy term, where s and t represent source and sink, respectively, and v and v' represent the two pixels involved in the energy term. In this case $E(1,0) - E(0,0) > 0$ and $E(1,0) - E(1,1) > 0$. (b) The graph obtained at the end results from adding elementary graphs.

The structure of the graph is as follows: first build vertices and edges corresponding to each pair of neighboring pixels, and then join these graphs together based on the additivity theorem also given in [38].

So, for each energy term E_h^{ij} and E_v^{ij} [see expression (15)], we construct an "elementary" graph with four vertices $\{s, t, v, v'\}$, where $\{s,t\}$ represents source and the sink, common to all terms, and $\{v, v'\}$ represents the two pixels involved [v being the left (up) pixel and v' the right (down) pixel]. Following very closely [38], we define a directed edge (v, v') with the weight $E(0,1) + E(1,0) - E(0,0) - E(1,1)$. Moreover, if $E(1,0) - E(0,0) > 0$, we define an edge (s, v) with the weight $E(1,0) - E(0,0)$ or, otherwise, we define an edge (v, t) with the weight $E(0,0) - E(1,0)$. In a similar way for vertex v', if $E(1,1) - E(1,0) > 0$, we define an edge (s, v') with weight $E(1,1) - E(1,0) > 0$ or, otherwise, we define an edge (v', t) with the weight $E(1,0) - E(1,1)$. Fig. 2(a) shows an example where $E(1,0) - E(0,0) > 0$ and $E(1,0) - E(1,1) > 0$. Fig. 2(b) illustrates the complete graph obtained at the end.

D. Energy Minimization Algorithm

Algorithm 1 PUMA: Graph Cuts Based Phase Unwrapping Algorithm

Initialization $\mathbf{k} \equiv \mathbf{k}' \equiv \mathbf{0}$, possible_improvement $\equiv 1$

while possible-improvement **do**
  Compute $E(0,0), E(1,1), E(0,1)$, and $E(1,0)$ {for every horizontal and vertical pixel pair}.
  Construct elementary graphs and merge them to obtain the main graph.
  Compute the max-flow/min-cut (S,T) {S-source set; T-sink set}.
  for all pixel (i,j) **do**
    if pixel $(i,j) \in S$ **then**
      $\boldsymbol{k'}_{i,j} = \boldsymbol{k}_{i,j} + 1$
    else
      $\boldsymbol{k'}_{i,j} = \boldsymbol{k}_{i,j}$ {remains unchanged}
    end if

11: **end for**
12: **if** $E(\mathbf{k}'|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi})$ **then**
13: $\mathbf{k} = \mathbf{k}'$
14: **else**
15: possible-improvement $= 0$
16: **end if**
17: **end while**

Algorithm 1 shows the pseudo-code for the phase unwrapping max-flow (PUMA) algorithm. It solves a sequence of binary optimizations until no energy decreasing is possible.

Concerning computational complexity, PUMA takes $N_{\text{bopt}} \times N_{mf}$ flops (measured in number of floating point operations), where N_{bopt} and N_{mf} stand for number of binary optimizations and number of flops per max-flow computation, respectively. In Section III-B, we have proofed that the algorithm stops in K iterations, where K is the range of E in wrap counts. Therefore, $N_{\text{bopt}} = K$. Concerning N_{mf}, in the results presented in Section V, we have used the augmenting path type max-flow/min-cut algorithm proposed in [41]. The worst-case complexity for augmenting path algorithms is $O(n^2m)$ [58], where n and m are the number of vertices and edges, respectively. However, in a huge array of experiments conducted in [41], authors systematically found out a complexity that is inferior to that of the push-relabel algorithm [59], with the queue-based selection rule, which is $O(n^2\sqrt{m})$. Thus, we herein take this bound.

Given that in our graphs $m \simeq 3n$ and N_{bopt} does not depend on n, the worst case complexity of the PUMA algorithm is bounded above by $O(n^{2.5})$. In Section V, we run a set of experiments where the worst case complexity is roughly $O(n)$. This scenario has systematically been observed.

Regarding memory usage, PUMA requires $7n$ bytes.

E. Clique Potentials

So far, we have assumed the clique potentials to be convex. This is central in the two main results in the paper: the Theorem 1 and the regularity of energy (2). Both are implied by the inequality (34)

$$V(a) + V(c) - V(b) \geq V(a + c - b) \tag{19}$$

shown in Appendix, where $\min(a, c) \leq b \leq \max(a, c)$.

What if we apply a function θ to the arguments of V? Using the notation $\theta(x) = x'$, we get the proposition

$$V(a') + V(c') - V(b') \geq V[(a + c - b)']. \tag{20}$$

Now, noting that, by construction,[4] a, b and c differ from each other by multiples of 2π, if we choose $\theta(x) = \mathcal{P}(x) + \alpha x$, where $\mathcal{P}$ is any 2π-periodic real valued function and $\alpha \in \mathbb{R}$, proposition (20) becomes

$$\begin{aligned} V(a') + V(c') - V(b') &\geq V[\mathcal{P}(a + c - b) + \alpha(a + c - b)] \\ &= V[\mathcal{P}(a) + \alpha(a + c - b)] \qquad (21) \\ &= V[(\mathcal{P}(a) + \alpha a) + (\mathcal{P}(a) + \alpha c) \\ &\quad - (\mathcal{P}(a) + \alpha b)] \qquad (22) \\ &= V(a' + c' - b'). \qquad (23) \end{aligned}$$

[4]Stated in the proof of Theorem 1.

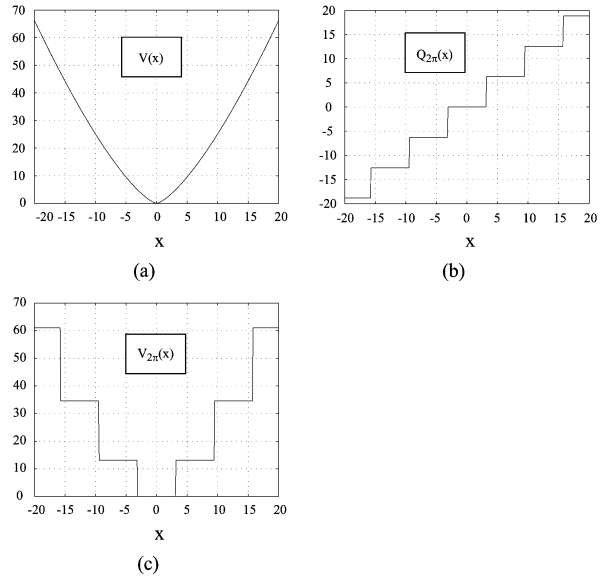

Fig. 3. (a) Convex function $C(x) = |x|^{1.4}$; (b) $Q_{2\pi}(x) = x - \mathcal{W}(x)$; (c) the classical $L^{1.4}$ norm potential given by $V_{2\pi}(x) = C[Q_{2\pi}(x)]$.

Since any 2π-sampling of θ is a monotone sequence, it is guaranteed that $\min(a', c') \leq b' \leq \max(a', c')$; so, proposition (23) follows from expression (19). Therefore, we have the following result.

Proposition 1: The set of clique potentials considered in Theorem 1 can be enlarged by admitting functions of the form $V \equiv C \circ (P + L)$, where C is a convex function, P is a 2π-periodic function, and L is a linear function.

It should be stressed that for such a potential, the regularity condition (18) is also satisfied; it follows directly from (23). We can, thus, conclude that the PUMA algorithm is valid for this broader class of clique potential functions. We next give some examples of possible clique potentials.

1) The Classical L^p Norm: By far, this is the most widely used class of clique potentials in phase unwrapping; it is given by $V(\Delta\phi) = |\Delta\phi - \mathcal{W}(\Delta\psi)|^p$, where $\mathcal{W}(x)$ is the principal phase value of x defined in the interval $[-\pi, \pi)$. In the jargon, $\mathcal{W}$ is termed the *wrapping operator*. Since $\Delta\phi$ and $\Delta\psi$ differ by a multiple of 2π, then $|\Delta\phi - \mathcal{W}(\Delta\psi)|^p = |\Delta\phi - \mathcal{W}(\Delta\phi)|^p$. Therefore, in our setting, we identify immediately $C(x) = |x|^p$, $\mathcal{P}(x) = -\mathcal{W}(x)$, and $A(x) = x$.

As stated in the Introduction, methods using this clique potential find a phase solution ϕ for which L^p norm of the difference between absolute phase differences and wrapped phase differences (so a second-order difference) is minimized.

From above, we see that C is convex given that $p \geq 1$. Therefore, we conclude that, for this range of p values, PUMA exactly solves the classical minimum L^p norm phase unwrapping problem.

From now on, we refer to $Q_{2\pi}(x) \equiv -\mathcal{W}(x) + x$ as the *2π-quantization function* and denote $V_{2\pi}(x) \equiv V[Q_{2\pi}(x)]$.

Fig. 3 plots the potential $C(x) = |x|^{1.4}$, the quantization function $Q_{2\pi}(x)$, and the classical $L^{1.4}$ norm given by $V_{2\pi}(x) = |Q_{2\pi}(x)|^{1.4}$.

2) Convex Potential: Choosing any convex $C(x)$, $P(x) = 0$ and $L(x) = x$, we obviously get back to the convex potential

case. For example, the quadratic clique potential $V(x) = x^2$ was used in work [7], under a Bayesian approach and a Markovian prior for the absolute phase. As already said, this potential tends to smooth phase discontinuities.

IV. Nonconvex Potentials

In image reconstruction, and in phase unwrapping in particular, images usually show a piecewise smooth spatial arrangement; this is a consequence of the smoothness of the imaged objects themselves, and of the discontinuities introduced by their borders. These discontinuities encode, then, relevant information that should be preserved in the reconstructed image.

It is well known that, in an energy minimization framework for image reconstruction, nonconvex clique potentials are desirable to allow discontinuity preservation (see, e.g., [60, Ch.3]) for discussion about discontinuity adaptive potentials). We should note here that, as we have shown in Section III-E, formally, a nonconvex clique potential is allowed in the algorithm, as long as every 2π-periodic sampling is convex (about the issue of convex functions on discrete domains see, e.g., [61]). It is, however, a trivial reasoning to conclude that this kind of nonconvex potentials are not discontinuity preserving. We will not enter, in this paper, into further detail on this subject.

A general nonconvex potential, nevertheless, makes the above introduced algorithm not valid and the reason is twofold. First, Theorem 1 demands a 2π-periodically convex V, i.e., a potential V such that every 2π-periodic sampling of it is convex. Let us use the terminology of [39] and call a 1-jump move the operation of adding a binary image $\boldsymbol{\delta}$; so, if V is nonconvex it is not possible, in general, to reach the minimum through 1-jump moves only. Second, as we emphasize in the sequence, it is trivial to show that, with a general nonconvex V, condition (18) does not hold with generality for every horizontal and vertical pairwise clique interaction. This means that we *cannot* apply the energy graph-representation used in the binary optimization employed in Algorithm 1 in Section III-D.

We now devise an approximate algorithm as a minor modification of PUMA to handle those two issues.

Regarding the latter, as the problem relies on the nonregularity of some energy terms $E^{ij}(\delta_i, \delta_j)$, i.e., they do not verify (18), our procedure consists in approximating them by regular ones. We do that by leaning on majorize minimize MM [47] concepts. Assume that we still want to minimize $E(\mathbf{k}^t + \boldsymbol{\delta}|\boldsymbol{\psi})$ given by (16). $E(\mathbf{k}^t|\boldsymbol{\psi})$ corresponds to $\boldsymbol{\delta} = 0$ and, therefore, to $\delta_i = 0$. Consider the regular energy $E^{'ij}(\delta_i, \delta_j)$ such that

$$\begin{cases} E^{'ij}(\delta_i, \delta_j) \geq E^{ij}(\delta_i, \delta_j), & \text{if } (\delta_i, \delta_j) \neq (0,0) \\ E^{'ij}(0,0) = E^{ij}(0,0), & \text{if } \quad (\delta_i, \delta_j) = (0,0) \end{cases} \tag{24}$$

i.e., $E^{'ij}$ majorizes E^{ij}. Define $Q(\boldsymbol{\delta}) = \sum_{ij \in \mathbb{Z}_1} E^{'ij}(\delta_i, \delta_j)$ and $\boldsymbol{\delta}^* = \min_{\boldsymbol{\delta}} Q(\boldsymbol{\delta})$. Then

$$E(\mathbf{k}^t + \boldsymbol{\delta}^*|\boldsymbol{\psi}) \leq Q(\boldsymbol{\delta}^*) \leq Q(\mathbf{0}) = E(\mathbf{k}^t|\boldsymbol{\psi}).$$

Therefore, the sequence $\{E(\mathbf{k}^t|\boldsymbol{\psi}), t = 0, 1, \ldots\}$ is decreasing.

A possible solution to obtain the replacement terms is, for instance, to increase term $E^{ij}(0,1)$ until $[E^{ij}(0,1) + E^{ij}(1,0) - E^{ij}(0,0) - E^{ij}(1,1)]$ equals zero; the corresponding graph of the Fig. 2 has no more negative edge weights. This solution, while may not be the best (concerning energy decreasing), is the simplest to implement: by observing that $E^{ij}(0,1)$ does not enter into any of the source/sink edges in the graph, it suffices to set the (v, v') interpixel edge (see Section III-C) weight to zero (thus assuring regularity).

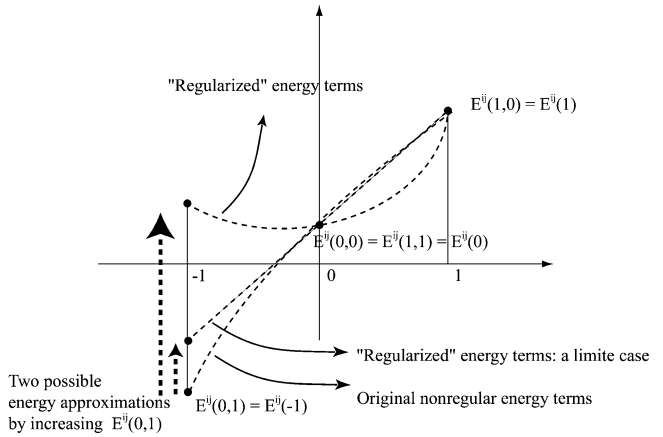

Fig. 4. Replacing nonregular energy terms by regular ones; we end up with an approximate energy. One of the possible approximations is to increase $E^{ij}(0,1)$.

In Fig. 4, we illustrate this energy approximation. We recall that, using a notation abuse, $E^{ij}(\delta_i, \delta_j) = E^{ij}(\delta_i - \delta_j)$ [see (2) and (15)]. The regularity condition (18), thus, can be written as

$$E^{ij}(0) \leq \frac{E^{ij}(-1) + E^{ij}(1)}{2} \tag{25}$$

which, being a convexity expression, means that regularity and convexity are equivalent for the energy that we are considering. Continuous convex and concave functions are shown to emphasize the regular/convex and nonregular/nonconvex parallel.[5] We note again that other energy approximations are possible and eventually even better; for instance, equally increasing $E^{ij}(0,1)$ and $E^{ij}(1,0)$ until condition (25) is satisfied. This issue is, however, out of the scope of this paper.

With respect to the first referred reason for non validity of PUMA, our strategy is to extend the range of allowed moves. Instead of only 1-jumps we now use sequences of s-jumps, introduced in [39], which correspond to add an $s\boldsymbol{\delta}$ image (increments can have 0 or s values).

The approximate algorithm presented has proved outperforming results in all the experiments we have put it through; in the next section, we illustrate some of that experiments. Algorithm 2 shows its pseudo-code.[6]

Algorithm 2 PUMA (Nonconvex Cliques)

Initialization: $\mathbf{k} \equiv \mathbf{k}' \equiv \mathbf{0}$

1: **for all** s in $[1, 2 \ldots, m, 1, 2, \ldots, m]$ (m is the maximum jump size) **do**
2: possible-improvement $\equiv 1$
3: **while** possible-improvement **do**

[5]It should be noted that discrete functions $f : \mathbb{Z} \longrightarrow \mathbb{R}$ are convex *iff* there exists an extension of f, $\overline{f} : \mathbb{R} \longrightarrow \mathbb{R}$, that is also convex.

[6]We note that, preferably, the maximum jump size should be chosen to be equal to the range of values of the unwrapped surface divided by 2π.

Compute $E(0,0), E(1,1), E(0,1)$, and $E(1,0)$ {for every horizontal and vertical pixel pair}.
Find nonregular pixel pairs $[E(0,1)+E(1,0)-E(0,0)-E(1,1)<0]$. If there is any, regularize it using the MM method (for instance, set the linking edge weight to zero).
Construct elementary graphs and merge them to obtain the main graph.
Compute the max-flow/min-cut (S,T) {S-source set; T-sink set}.
for all pixel (i,j) **do**
  if pixel $(i,j) \in S$ **then**
    $\boldsymbol{k}'_{i,j} = \boldsymbol{k}_{i,j} + s$
  else
    $\boldsymbol{k}'_{i,j} = \boldsymbol{k}_{i,j}$ {remains unchanged}
  end if
end for
if $E(\mathbf{k}'|\boldsymbol{\psi}) < E(\mathbf{k}|\boldsymbol{\psi})$ **then**
  $\mathbf{k} = \mathbf{k}'$
else
  possible-improvement $= 0$
end if
end while
end for

Finally, it should be noted that the question of what particular nonconvex potential to choose is a relevant one. The main problems, in phase unwrapping, arise both from noise and from discontinuities presence. The small amplitude noise (variance smaller than π) is well described by a Gaussian density, meaning that the potentials near the origin should be quadratic. In what relates to larger amplitude discontinuities, they should not be too much penalized and, as such, it makes sense to employ potentials growing much slower than the quadratic. This is why it makes sense to choose potentials like, e.g., the truncated quadratic [43] and the potential used by Geman and Mclure [62].

V. PUMA Application Examples

In this section, we briefly illustrate PUMA performance on representative phase unwrapping problems. The results presented were obtained with MATLAB coding (max-flow algorithm is implemented in C++),[7] and using a PC workstation equipped with a 1.7-Ghz Pentium-IV CPU.

Fig. 5(a) and (b) displays two phase images (256×256 pixels) to be unwrapped; they are synthesized from original absolute phase surfaces formed by Gaussian elevations with heights of 25π and 50π rad, respectively, and common standard deviations $\sigma_i = 25$ and $\sigma_j = 40$ pixels, in the vertical and horizontal dimensions, respectively. The wrapped images are generated according to an InSAR observation statistics (see, e.g., [7]), producing an interferometric pair, with correlation coefficient 0.7 and 1.0, respectively. The wrapped phase images are, then, obtained (for each pair), by computing the product of one image by the complex conjugate of the other, and finally taking the argument.

[7]Max-flow code made available at http://www.cs.cornell.edu/People/vnk/software.html by V. Kolmogorov. See [41] for more details.

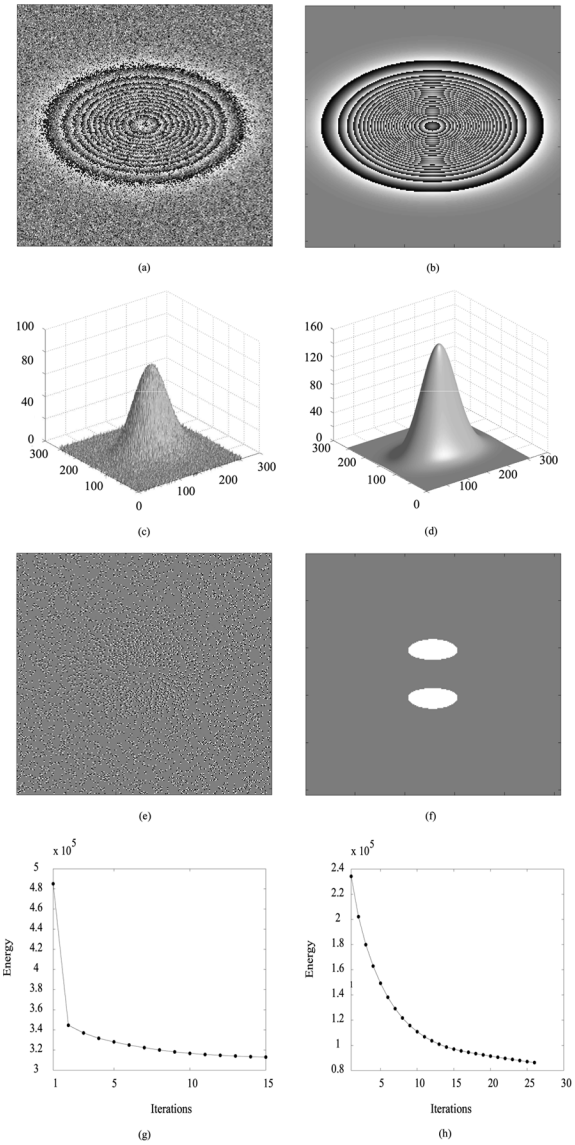

Fig. 5. (a) Wrapped Gaussian elevation with 25π height. The associated noise standard deviation is 1.07 rad. (b) Wrapped Gaussian elevation with 50π height. The associated noise standard deviation is 0 rad. (c) Image in (a) unwrapped by PUMA. (d) Image in (b) unwrapped by PUMA. (e) Residues on the image presented in (a): white and black pixels means positive and negative residues, respectively. (f) Aliased regions (signalled by white pixels) of the image in (b). (g) Energy decreasing for the unwrapping of image in (a). (h) Energy decreasing for the unwrapping of image in (b).

Regarding the first image [Fig. 5(a)], the coherence value of 0.7 corresponds to a noise whose standard deviation is 1.07 rad, thus inducing a huge number of phase jumps (*residues*), making the unwrapping a hard task. Fig. 5(c) shows the corresponding unwrapped surface by PUMA using a nonquantized L^2 norm potential. Even with low-correlation induced discontinuities, PUMA successfully accomplishes a correct unwrapping (error free). We emphasize that our algorithm seeks the correct wrap-count image, so it does not intend to get rid of the possible existing noise, whatsoever. Regarding the second image [Fig. 5(b)], although the coherence value is at the maximum (there is no noise), it presents phase rates large enough to produce aliasing, such that the unwrapping becomes a hard task. Fig. 5(d) shows the corresponding unwrapped surface by

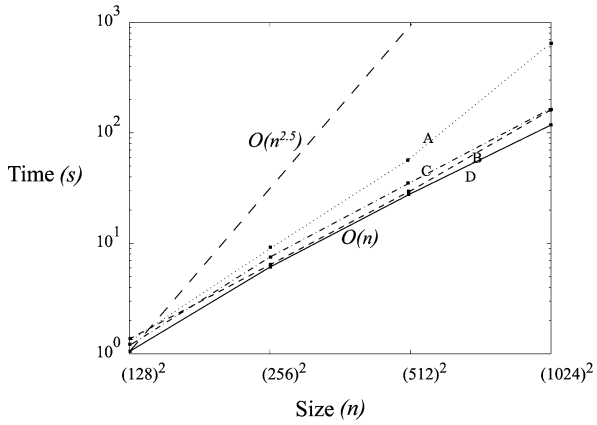

Fig. 6. Unwrapping times of a 14π height Gaussian surface with PUMA, using a PC workstation equipped with a 1.7-Ghz Pentium-IV CPU: time (s) versus image size (n). Time grows roughly as $O(n)$ in all the four shown experiments. An $O(n^{2.5})$ line is shown for reference. (A) Gaussian surface with 1.07 rad interferometric noise unwrapped with a nonquantized L^2 norm. (B) Gaussian surface without interferometric noise unwrapped with a nonquantized L^2 norm. (C) Gaussian surface with 1.07 rad interferometric noise unwrapped with a classical (quantized) L^2 norm. (D) Gaussian surface without interferometric noise unwrapped with a classical (quantized) L^2 norm.

PUMA using again a nonquantized L^2 norm potential. Even with aliasing induced discontinuities, PUMA successfully accomplishes a correct unwrapping (error free). For both the unwrappings, we have chosen the nonquantized L^2 norm potential, as it shows a good performance regarding the unwrapping of this kind of noisy/aliased wrapped surfaces [7]. Fig. 5(e) shows the residues existing on the image shown in Fig. 5(a); white pixels are positive residues and black pixels are negative residues. We point out that it was not supplied any discontinuity information to the algorithm. Fig. 5(f) shows the regions of the original image that present aliasing (white pixels region). Figs. 5(e) and 5(f) show the energy evolution along the fifteen and 26 iterations taken by the algorithm to perform the unwrapping of the images in Figs. 5(a) and 5(b), respectively. It is noticeable a major energy decreasing in the first few iterations.

As referred to in Section III-D, we have observed approximately an $O(n)$ complexity (where n is the size of the input image) in the experiences we have run with PUMA. Fig. 6 illustrates this for the unwrapping of the Gaussian surface with and without noise, and employing two kinds of clique potentials.

Fig. 7(a) is analogous to Fig. 5(a), but now the original phase surface is a Gaussian with a 20π rad height and a quarter of the plane set to zero. This null quarter causes, therefore, many discontinuities, which renders a very difficult phase unwrapping problem. It should be noted that, again, we do not provide any discontinuity information to PUMA in this experiment. Fig. 7(b) shows the tentative unwrapped image with a classical L^1 norm. With such a potential, the computed phase is useless. Fig. 7(c) shows a successful, with an error of $3\times 2\pi$ in just one pixel (the dark among white ones in the border), unwrapping in 12 iterations, for which the energy decreasing is shown in Fig. 7(h). Fig. 7(d) shows the mesh corresponding to 7(c). This unwrapping was obtained using the approximate version of PUMA with the nonconvex potential depicted in Fig. 7(g), and a maximum jump size $m = 1$. In Fig. 7(e) and 7(f), we show, respectively, the nonregular horizontal and vertical cliques during the first iteration

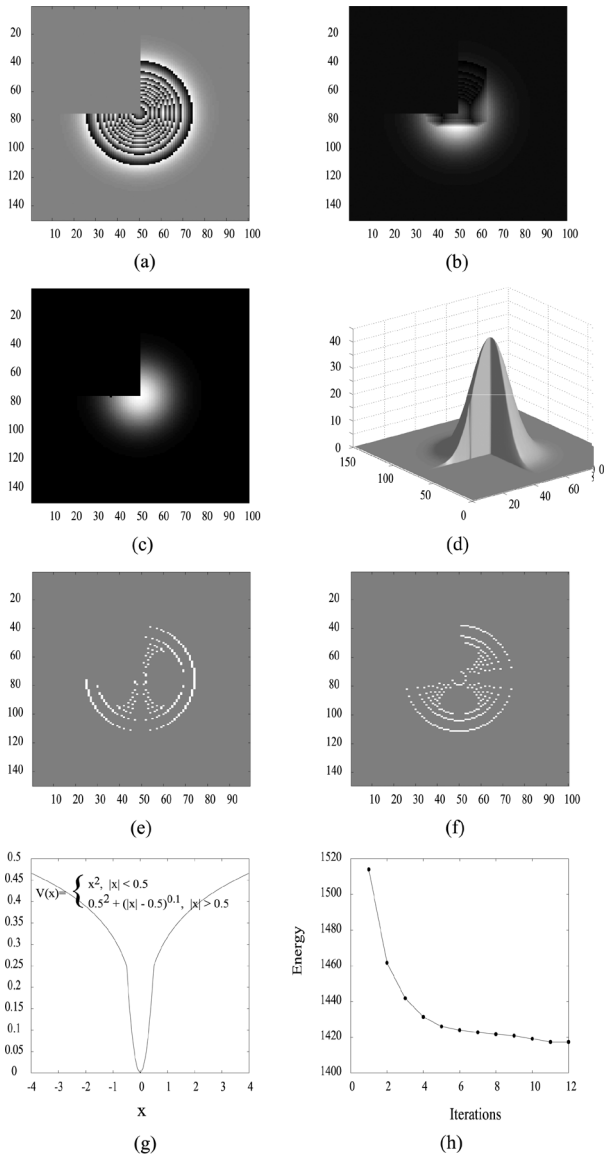

Fig. 7. (a) Wrapped Gaussian elevation with a quarter of the plane with zero height. (b) Image in (a) tentatively unwrapped with a classical L^1 norm clique potential. (c) Image in (a) successfully unwrapped ($3\times 2\pi$ error in one pixel) using a nonconvex clique potential. (d) A "3-D" rendering of the unwrapped image. (e) Nonregular horizontal cliques (white signalled) during the first iteration (successful unwrapping). (f) Nonregular vertical cliques (white signalled) during the first iteration (successful unwrapping). (g) Nonregular clique potential employed. (h) Energy decreasing along the successful unwrapping.

of the algorithm (signalled as white). The number of nonregular cliques is relatively small (235 and 243, respectively).

Fig. 8(a) shows a phase image (152×458 pixels) to be unwrapped. It was obtained from an original absolute phase surface, that corresponds to a (simulated) InSAR acquisition for a real steep-relief mountainous area inducing, therefore, many discontinuities and posing a very tough PU problem. This area corresponds to Long's Peak, CO, and the data is distributed with book [5]. The wrapped image is generated according to an InSAR observation statistics (see, e.g., [14]), producing an interferometric pair; by computing the product of one image of the pair by the complex conjugate of the other and finally taking the argument, the wrapped phase image is then obtained.

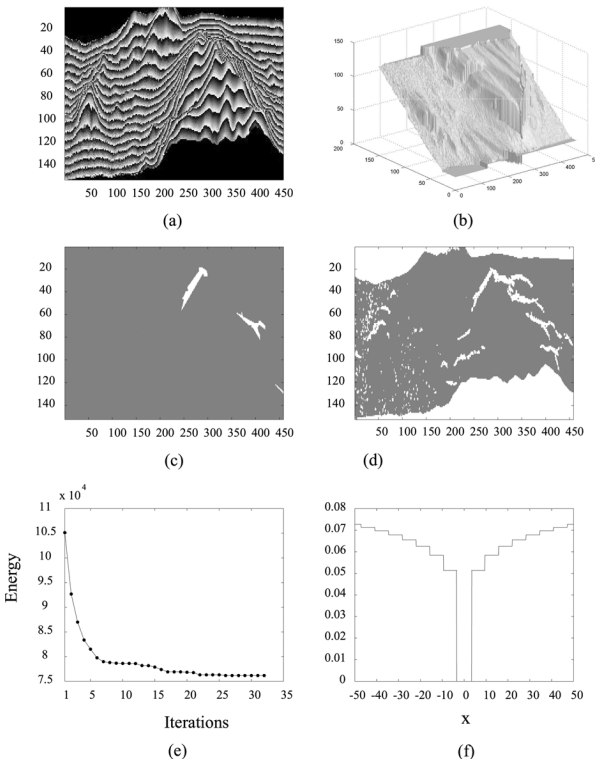

Fig. 8. (a) Wrapped phase image obtained from a simulated InSAR acquisition from Long's Peak, CO (data distributed with [5]). (b) Image in (a) unwrapped by PUMA (32 iterations). (c) Discontinuity information given as input to the unwrapping process. White pixels signal discontinuity locations. (d) The total discontinuity information at disposal. White pixels signal discontinuity locations. (e) Energy decreasing for the unwrapping of image in (a). (f) The potential employed.

Fig. 8(d) shows a quality map (also distributed with book [5]) computed from the InSAR coherence estimate (see [5, Ch.3]) for further details). However, to illustrate the discontinuity preserving ability of the PUMA method with nonconvex potentials, we have reduced, substantially, the number of supplied discontinuities in the algorithm. The corresponding quality map is shown in Fig. 8(c). The PU problem thus obtained is far more difficult than the original (i.e., using the complete quality map) and a nonconvex potential is able to solve it. The resulting phase unwrapped is "3-D" rendered in Fig. 8(b), corresponding to an error norm (variance of the image given by the difference between original and unwrapped phase images) of 0.6 rad^2. The unwrapping was obtained using the approximate version of PUMA, with $m = 2$. In Fig. 8(f), the employed nonconvex, quantized, potential is depicted. The correspondent analytical expression is given by $V_{2\pi}(x) = [Q_{2\pi}(x)]^{0.002}$. Fig. 8(e) illustrates the energy evolution with the algorithm iterations.

Fig. 9(a) shows another phase image (257×257 pixels) to be unwrapped, which was synthesized from an original surface (distributed with the book [5]) consisting of two "intertwined" spirals built on two sheared planes. It should be noticed that the original phase surface has many discontinuities, which make this an extremely difficult unwrapping problem, if no information is supplied about discontinuities locations. The approximate version of PUMA is able to blindly unwrap this image as is shown in

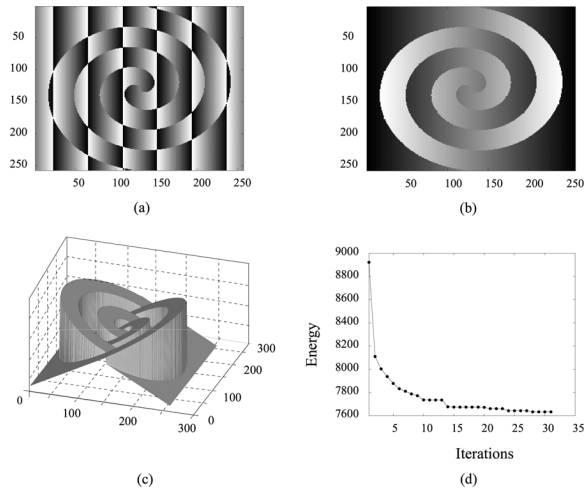

Fig. 9. (a) Wrapped phase image corresponding to an original phase surface of two intertwined spirals in two sheared planes (data distributed with [5]). (b) Image in (a) blindly unwrapped by PUMA (31 iterations). (c) A "3-D" rendering of the unwrapped image. (d) Energy decreasing for the unwrapping of image in (a). Notice that no discontinuities are supplied to the algorithm.

Fig. 9(b), by using a maximum jump size $m = 7$ and a nonconvex potential given by the following analytical expression:

$$V(x) = \begin{cases} 0.5^{(0.001-2)}x^2, & |x| \le 0.5 \\ |x|^{0.001}, & |x| > 0.5. \end{cases} \tag{26}$$

Fig. 9(c) shows a "3-D" rendering of the unwrapped surface and Fig. 9(d) shows the decreasing of the energy, along 31 iterations, in the unwrapping process.

We emphasize that we obtained a correct (error free) unwrapping except for a few (ten or so) pixels; these are pixels that in image Fig. 9(a) are in the border of the two spirals and furthermore present continuity with both vertical and horizontal neighbors. This is considered an image artifact and not an error of the algorithm.

Fig. 10(a) shows another phase image (256×256 pixels) to be unwrapped. As in [31], it is a kind of cylinder upon a ramp and has a uniform noise of 3 rad. The result of unwrapping this image using the approximate version of PUMA is shown in Fig. 10(b). It was employed the nonconvex potential

$$V(x) = \begin{cases} 2^{(0.01-2)}x^2, & |x| \le 2 \\ |x|^{0.01}, & |x| > 2 \end{cases} \tag{27}$$

and a maximum jump size of $m = 9$. Fig. 10(c) shows the pixels where the unwrapping went wrong; it amounts to only 0.39% of the total pixels. It should be noticed that no discontinuity information was supplied to the algorithm, which employed 43 iterations along nearly 100 s. Fig. 10(d) depicts the employed potential. The results here presented show an apparent more accurate and fast phase unwrapping than those reported in [31] (note that we use four neighbors for each pixel).

Given the quite different problems presented above, a natural question arises: Given a certain particular PU problem, what potentials and scheduling should be taken? Firstly, we stress that a proper answer to this question is out of the scope of the present paper. However, let us remark that for the problem in Fig. 5 any kind of nonquantized potential proved to work in all the experiences we have put it through; regarding Fig. 7, we found that most

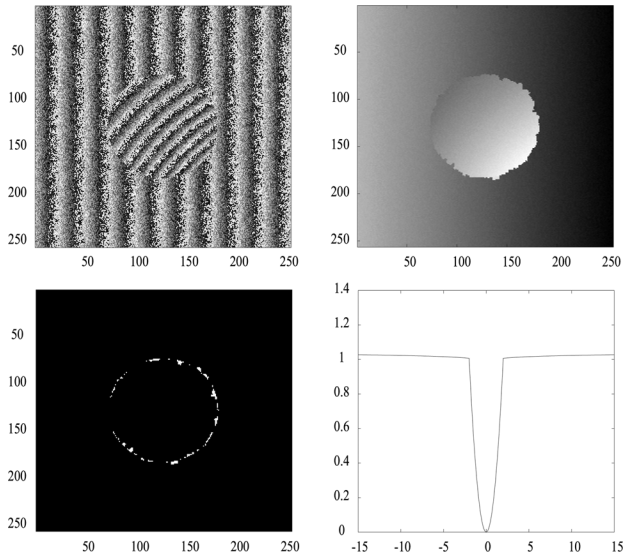

Fig. 10. (a) Wrapped phase image corresponding to an original phase surface given by a kind of cylinder upon a ramp. In all the images, there is a uniform noise of 3 rad (data reported in [31]). (b) Image in (a) blindly unwrapped by PUMA (43 iterations). (c) 0.39% of the total number of pixels (shown in white) had a wrong unwrapping. (d) The potential employed.

of the potentials with exponent less than one are suitable; with respect to Figs. 8 and 10, a more fine tuning of an appropriate potential had to be carried out, and finally, to attain a successfully solution of the difficult problem presented in Fig. 9, an even more thorough fine tuning was required. Regarding scheduling, a similar pattern of tuning needs was found; the values presented in the paper proved to give small schedules good enough to solve the problems. Although we put emphasis in that PUMA algorithm allows a large family of potentials and move spaces, which gives flexibility to address different phase unwrapping problems, the question of finding the more suitable ones is relevant and to be addressed in the future. It is worth mentioning that there exist regularization criteria that are, perhaps, more suitable at handling PU problems due to aliasing induced by high gradients, or due to noisy images. As an example, we mention work [63] which employs a second-order Markov random field. Nevertheless, in what regards blindly preserving discontinuities, the criteria that we adopt here perform very well.

VI. Concluding Remarks

We developed a new graph cuts based phase unwrapping methodology, which embraces, in particular, the minimum L^p norm class of PU problems. The iterative binary optimization sequence proposed in the $\mathbb{Z}\pi$M algorithm [7] was generalized to a broader family of clique potentials; this broader set is now given by the composition of any convex pairwise function, depending only on differences, with the sum of a 2π-periodic with a linear function. Furthermore, each binary optimization problem in the above referred sequence is solved by applying results on energy minimization using graph cuts from [38]. These optimizations are computed efficiently using max-flow/min-cut algorithms well known from combinatorial optimization. The proposed algorithm, termed PUMA, is an exact solver, in particular, of the minimum L^p norm class of PU algorithms, for $p \geq 1$; its computational complexity is $KT(n, 3n)$, where K is the number of 2π multiples and $T(n, m)$ is the complexity of a max-flow computation in a graph with n nodes and m edges. In practice, we have observed that the complexity is $O(n)$, what is in line with other reports on graph cuts based optimization.

Moreover, we have also addressed the phase discontinuities issue by employing nonconvex discontinuity preserving potentials. As this turns out to be an NP-hard problem, we devised an approximate version of the PUMA algorithm by leaning on two main ideas: first, to apply majorize-minimize approximation, which allows us to exploit the graph cuts binary optimization framework; second, to enlarge the size of allowable binary moves, thus coping with the local minima arising from nonconvex potentials. A set of phase unwrapping experiments presented illustrates the state-of-the-art discontinuity preserving abilities of PUMA.

Appendix

Proof of Theorem 1: This proof parallels the proofs of Lemma 1 in the appendixes of [7] and of [19], with the appropriate modifications to deal with the more general clique potentials here employed.

Define $\Delta k_{ij} = [k_2]_{ij} - [k_1]_{ij}$, for $(i,j) \in \mathbb{G}_0$ where $\mathbb{G}_0 = \{(i,j) : i = 1, \ldots, M, j = 1, \ldots, N\}$, with M and N denoting the number of lines and columns respectively (i.e., the usual image pixel indexing 2-D grid). Given that the energy $E(\mathbf{k}|\boldsymbol{\psi})$ depends only on differences between elements of $\mathbf{k}$, we take $\Delta k_{ij} \geq 0$ for $(i,j) \in \mathbb{G}_0$. Define $n = \max_{ij}(\Delta k_{ij})$ and the wrap-count image sequence $\{\mathbf{k}^{(t)}, t = 0, \ldots, n\}$, such that $\mathbf{k}^{(0)} = \mathbf{k}_1$, $\mathbf{k}^{(n)} = \mathbf{k}_2$, and

$$k_{ij}^{(t)} = k_{ij}^{(0)} + \min(t, \Delta k_{ij}), \quad t = 0, \ldots, n. \tag{28}$$

The energy variation $\Delta E \equiv E(\mathbf{k}_2|\boldsymbol{\psi}) - E(\mathbf{k}_1|\boldsymbol{\psi})$ can be decomposed as

$$\Delta E \equiv \sum_{t=1}^{n} \underbrace{\left[E(\mathbf{k}^{(t)} \Big| \boldsymbol{\psi}) - E(\mathbf{k}^{(t-1)} \Big| \boldsymbol{\psi})\right]}_{\Delta E^{(t)}}.$$

Since $\Delta E < 0$ by hypothesis, then at least one of the terms $\Delta E^{(t)}$ of the above sum is negative. The theorem is proved if we show that the variation $\delta E^{(t)} \equiv E(\mathbf{k}^{(0)} + \boldsymbol{\delta}^{(t)}|\boldsymbol{\psi}) - E(\mathbf{k}^{(0)}|\boldsymbol{\psi})$ satisfies $\delta E^{(t)} \leq \Delta E^{(t)}$, where $\boldsymbol{\delta}^{(t)} \equiv \mathbf{k}^{(t)} - \mathbf{k}^{(t-1)}$, for any $t = 1, \ldots, n$. This condition is equivalent to

$$\begin{aligned} 0 \leq\ & E(\mathbf{k}^{(t)}|\boldsymbol{\psi}) - E(\mathbf{k}^{(t-1)}|\boldsymbol{\psi}) \\ & - E(\mathbf{k}^{0} + \mathbf{k}^{(t)} - \mathbf{k}^{(t-1)}|\boldsymbol{\psi}) + E(\mathbf{k}^{0}|\boldsymbol{\psi}) \end{aligned} \tag{29}$$

for $t = 1, \ldots, n$. Introducing (2) into (29), we obtain $0 \leq S^h + S^v$, where

$$\begin{aligned} S^h = \sum_{ij} \Big[& V\left(\Delta\phi_{ij}^{h(t)}\right) - V\left(\Delta\phi_{ij}^{h(t-1)}\right) + V\left(\Delta\phi_{ij}^{h(0)}\right) \\ & - V\left(\Delta\phi_{ij}^{h(0)} + \Delta\phi_{ij}^{h(t)} - \Delta\phi_{ij}^{h(t-1)}\right)\Big] \overline{v}_{ij} \end{aligned} \tag{30}$$

$$\begin{aligned} S^v = \sum_{ij} \Big[& V\left(\Delta\phi_{ij}^{v(t)}\right) - V\left(\Delta\phi_{ij}^{v(t-1)}\right) + V\left(\Delta\phi_{ij}^{v(0)}\right) \\ & - V\left(\Delta\phi_{ij}^{v(0)} + \Delta\phi_{ij}^{v(t)} - \Delta\phi_{ij}^{v(t-1)}\right)\Big] \overline{h}_{ij} \end{aligned} \tag{31}$$

where V is the clique potential, and $\Delta\phi_{ij}^{h(t)}$ and $\Delta\phi_{ij}^{v(t)}$ are given by (3) and (4), respectively, computed at the wrap-count image $\mathbf{k}^{(t)}$. To prove (29), we now show that the terms of S^h corresponding to a given site $(i,j) \in \mathbb{G}_1$ have positive sum. The same is true concerning S^v.

The difference $k_{ij}^{(t)} - k_{ij-1}^{(t)}$, for $t = 0,\ldots,n$, is a monotone sequence. This is a consequence of the definition (28): if $\Delta k_{ij} > \Delta k_{ij-1}$ the sequence is monotone increasing; if $\Delta k_{ij} \leq \Delta k_{ij-1}$ the sequence is monotone decreasing. Therefore, the sequence $\{\Delta\phi_{ij}^{h(t)}\}$, for $t = 0,\ldots,n$, is also monotone. Define $a \equiv \Delta\phi_{ij}^{h(0)}$, $b \equiv \Delta\phi_{ij}^{h(t-1)}$, and $c \equiv \Delta\phi_{ij}^{h(t)}$, and without loss of generality let us assume[8] $a \geq b \geq c$. We will show that the sum of terms of S^h, corresponding to the site (i,j) is positive

$$\begin{aligned} V(c) - V(b) + V(a) - V(a+c-b) \geq 0 \\ V(a) + V(c) - V(b) \geq V(a+c-b). \end{aligned} \tag{32}$$

By hypothesis, V is convex. Also by hypothesis, $a \geq b \geq c$, so $\exists t \in [0,1] : b = at + c(1-t)$. Thus

$$\begin{aligned} V(b) &\leq tV(a) + (1-t)V(c) \\ V(a) + V(c) - V(b) &\geq V(a) + V(c) \\ &\quad - [tV(a) + (1-t)V(c)] \\ &\geq (1-t)V(a) + tV(c). \end{aligned} \tag{33}$$

As V is convex, $(1-t)V(a) + tV(c) \geq V[(1-t)a + tc]$. So, from (33)

$$\begin{aligned} V(a) + V(c) - V(b) &\geq V[(1-t)a + tc] \\ &\geq V(a + c - \underbrace{[at + c(1-t)]}_{b}) \\ &\geq V(a+c-b). \end{aligned} \tag{34}$$

The same reasoning applies to S^v. ■

Proof of Lemma 1: The proof is inspired in the Proposition 3.7 of [48]. The main difference is that the class of energies herein considered does not have unary terms. The implication of this is that our steepest descent algorithm, in each steep, finds a move in the set $\mathcal{B} = \{0,1\}^{\mathrm{MN}}$, whereas the presence of unary terms imposes the search in the larger set $\{-1,0,1\}^{\mathrm{MN}}$, as proposed in [48, Ch.3.3] and [49].

Define $\mathbf{u} = \mathbf{k}^t$ and $E(\cdot) \equiv E(\cdot\,|\,\boldsymbol{\psi})$. Let $\mathcal{M}_{t+1}$ be the set of minimizers of $E(\cdot)$ on $\mathcal{D}_{t+1}$. If $E(\mathbf{v}) = E(\mathbf{u})$ for $\mathbf{v} \in \mathcal{M}_{t+1}$, then $\mathbf{u} \in \mathcal{M}_{t+1}$ and the lemma is proved by choosing $\boldsymbol{\delta} = \mathbf{0}$. Let us then assume that $E(\mathbf{v}) < E(\mathbf{u})$ for $\mathbf{v} \in \mathcal{M}_{t+1}$. We proceed by contradiction supposing that $\mathbf{v} - \mathbf{u} \notin \mathcal{B}$, for all $\mathbf{v} \in \mathcal{M}_{t+1}$, i.e., for all $\mathbf{v} \in \mathcal{M}_{t+1}$ there exists at least one site $i,j \in \mathbb{G}_0$ such that

$$v_{ij} - u_{ij} \notin \{0,1\}. \tag{35}$$

Given $\mathbf{u} \in \mathcal{M}_t$ and $\mathbf{v} \in \mathcal{M}_{t+1}$, define image $\mathbf{h}$ with $h_{ij} = 1$ if $v_{ij} - u_{ij} > 0$ and zero elsewhere. At least one element of $\mathbf{v}$ takes the value $t+1$ and all elements of $\mathbf{u}$ are less ou equal to t. Therefore, we have $\mathbf{h} \neq \mathbf{0}$.

[8]The only possibilities are either $a \geq b \geq c$ or $a \leq b \leq c$, because the sequence $\{\Delta\phi_{ij}^{h(t)}\}$ is monotone as we have shown.

Since $E(\cdot)$ is a linear combination of convex terms, each one depending only on a difference of two components, then a reasoning based on (32) leads to

$$E(\mathbf{u}) - E(\mathbf{v} - \mathbf{h}) \geq E(\mathbf{u} + \mathbf{h}) - E(\mathbf{v}).$$

The right-hand side of the above inequality is nonnegative, for $\mathbf{v}$ is a global minimizer in $\mathcal{D}_{t+1}$. If $E(\mathbf{u}+\mathbf{h}) = E(\mathbf{v})$, hypothesis (35) would be contradicted because $\mathbf{v} - \mathbf{u} \in \mathcal{B}$. We then have

$$E(\mathbf{u}) > E(\mathbf{v} - \mathbf{h})$$

but $\mathbf{v} - \mathbf{h} \in \mathcal{D}_t$. To verify this, let us analyze the differences $v_{ij} - h_{ij}$, having in mind that $h_{ij} \in \{0,1\}$ and $0 < v_{ij} \leq t+1$. If $v_{ij} = t+1$, then $h_{ij} = 1$ and $v_{ij} - h_{ij} = t$. Otherwise, $v_{ij} - h_{ij} \leq t$. Then, $\mathbf{v} - \mathbf{h} \in \mathcal{D}_t$, contradicting the fact that $\mathbf{u}$ is a global minimizer of E on $\mathcal{D}_t$. This ends the proof. ■

REFERENCES

[1] J. Bioucas-Dias and G. Valadão, "Phase unwrapping: A new max-flow/min-cut based approach," presented at the IEEE Int. Conf. Image Processing 2005.

[2] J. Bioucas-Dias and G. Valadão, , A. Rangarajan, B. Vemuri, and A. Yuille, Eds., "Discontinuity preserving phase unwrapping using graph cuts," in *Proc. Energy Minimization Methods in Computer Vision and Pattern Recognition*. New York: Springer, 2005, vol. 3757, pp. 268–284.

[3] L. Graham, "Synthetic interferometer radar for topographic mapping," *Proc. IEEE*, vol. 62, no. 2, pp. 763–768, Feb. 1974.

[4] H. Zebker and R. Goldstein, "Topographic mapping from interferometric synthetic aperture radar," *J. Geophys. Res. B*, vol. 91, no. 5, pp. 4993–4999, 1986.

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

[6] P. Rosen, S. Hensley, I. Joughin, F. LI, S. Madsen, E. Rodriguez, and R. Goldstein, "Synthetic aperture radar interferometry," *Proc. IEEE*, vol. 88, no. 3, pp. 333–382, Mar. 2000.

[7] J. Dias and J. Leitão, "The $\mathbb{Z}\pi$M algorithm for interferometric image reconstruction in SAR/SAS," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 408–422, Apr. 2002.

[8] L. Cutrona, "Comparison of sonar system performance achievable using synthetic aperture techniques with the performance achievable by more conventional means," *J. Acoust. Soc. Amer.*, vol. 58, no. 12, pp. 336–348, Aug. 1975.

[9] H. Griffiths, T. Rafik, Z. Meng, C. Cowan, H. Shafeeu, and D. Anthony, "Interferometic synthetic aperture sonar for high-resolution 3-D mapping of the seabed," *Proc. Inst. Elect. Eng. Radar, Sonar, Navigat.*, vol. 144, no. 2, pp. 96–103, Apr. 1997.

[10] P. Lauterbur, "Image formation by induced local interactions: Examples employing nuclear magnetic resonance," *Nature*, vol. 242, pp. 190–191, Mar. 1973.

[11] M. Hedley and D. Rosenfeld, "A new two-dimensional phase unwrapping algorithm for MRI images," *Magn. Res. Med.*, vol. 24, pp. 177–181, 1992.

[12] S. Pandit, N. Jordache, and G. Joshi, "Data-dependent systems methodology for noise-insensitive phase unwrapping in laser interferometric surface characterization," *J. Opt. Soc. Amer.*, vol. 11, no. 10, pp. 2584–2592, 1994.

[13] P. Jezzard and R. Balaban, "Correction for geometric distortion in echo-planar images from B_0 field variations," *Magn. Res. Med.*, vol. 34, pp. 65–73, 1995.

[14] B. Quesson, J. A. de Zwart, and C. T. Moonen, "Magnetic resonance temperature imaging for guidance of thermotherapy," *J. Magn. Res. Imag.*, vol. 12, no. 4, pp. 525–533, 2000.

[15] A. Rauscher, M. Barth, J. Reichenbach, R. Stollberger, and E. Moser, "Automated unwrapping of MR phase images applied to BOLD MR-venography at 3 Tesla," *J. Magn. Res. Imag.*, vol. 18, no. 2, pp. 175–180, 2003.

[16] K. Itoh, "Analysis of the phase unwrapping problem," *Appl. Opt.*, vol. 21, no. 14, 1982.

[17] R. Goldstein, H. Zebker, and C. Werner, "Satellite radar interferometry: Two-dimensional phase unwrapping," in *Proc. Symp. Ionospheric Effects on Communication and Related Systems*, 1988, vol. 23, pp. 713–720.

[18] J. Marroquin and M. Rivera, "Quadratic regularization functionals for phase unwrapping," *J. Opt. Soc. Amer.*, vol. 12, no. 11, pp. 2393–2400, 1995.
[19] T. Flynn, "Two-dimensional phase unwrapping with minimum weighted discontinuity," *J. Opt. Soc. Amer. A*, vol. 14, no. 10, pp. 2692–2701, 1997.
[20] M. Costantini, "A novel phase unwrapping method based on network programing," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 3, pp. 813–821, May 1998.
[21] M. Rivera, J. Marroquin, and R. Rodriguez-Vera, "Fast algorithm for integrating inconsistent gradient fields," *Appl. Opt.*, vol. 36, no. 32, pp. 8381–8390, 1995.
[22] 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, 1993.
[23] W. Xu and I. Cumming, "A region growing algorithm for insar phase unwrapping," in *Proc. Int. Geoscience and Remote Sensing Symp.*, Lincoln, NE, 1996, vol. 4, pp. 2044–2046.
[24] D. Ghiglia and L. 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.
[25] D. Ghiglia and L. Romero, "Minimum L^p norm two-dimensional phase unwrapping," *J. Opt. Soc. Amer.*, vol. 13, no. 10, pp. 1999–2013, 1996.
[26] C. Chen and H. Zebker, "Network approaches to two-dimensional phase unwrapping: Intractability and two new algorithms," *J. Opt. Soc. Amer.*, vol. 17, no. 3, pp. 401–414, 2000.
[27] J. Dias and J. Leitão, "Simultaneous phase unwrapping and speckle smoothing in SAR images: A stochastic nonlinear filtering approach," in *Proc. Eur. Conf. Synthetic Aperture Radar*, May 1998, pp. 373–377.
[28] G. Nico, G. Palubinskas, and M. Datcu, "Bayesian approach to phase unwrapping: Theoretical study," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2545–2556, Sep. 2000.
[29] C. Chen, "Statistical-cost network-flow approaches to two-dimensional phase unwrapping for radar interferometry," Ph.D. dissertation, Stanford Univ., Stanford, CT, 2001.
[30] O. Marklund, "An anisotropic evolution formulation applied in 2-D unwrapping of discontinuous phase surfaces.," *IEEE Trans. Image Process.*, vol. 10, no. 11, pp. 1700–1711, Nov. 2001.
[31] M. Rivera and J. Marroquin, "Half-quadratic cost functions for phase unwrapping," *Opt. Lett.*, vol. 29, no. 5, pp. 504–506, 2004.
[32] L. Ying, Z. Liang, D. Munson Jr., R. Koetter, and B. Frey, "Unwrapping of MR phase images using a markov random field model," *IEEE Trans. Med. Imag.*, vol. 25, no. 1, pp. 128–136, Jan. 2006.
[33] B. Friedlander and J. Francos, "Model based phase unwrapping of 2-d signals," *IEEE Trans. Signal Process.*, vol. 44, no. 12, pp. 2999–3007, Dec. 1996.
[34] Z. Liang, "A model-based method for phase unwrapping," *IEEE Trans. Med. Imag.*, vol. 15, no. 6, pp. 893–897, Jun. 1996.
[35] D. Fried, "Least-squares fitting a wave-front distortion estimate to an array of phase-difference measurements," *J. Opt. Soc. Amer.*, vol. 67, no. 3, pp. 370–375, 1977.
[36] J. Leitão 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.
[37] M. Datcu and G. Palubinskas, "Multiscale bayesian height estimation from insar using a fractal prior," *Proc. SPIE*, pp. 155–163, 1998.
[38] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 147–159, Feb. 2004.
[39] O. Veksler, "Efficient graph-based energy minimization methods in computer vision," Ph.D. dissertation, Cornell Univ., Ithaca, NY, 1999.
[40] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 11, pp. 1222–1239, Nov. 2001.
[41] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.
[42] A. Blake, "Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 1, pp. 2–12, Jan. 1989.
[43] A. Blake and A. Zisserman, *Visual Reconstruction*. Cambridge, MA: MIT Press, 1987.
[44] S. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 3, pp. 367–383, Mar. 1992.
[45] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. Med. Imag.*, vol. 8, no. 2, pp. 194–202, Jun. 1989.
[46] N. Vaidya and K. Boyer, "Discontinuity preserving surface reconstruction through global optimization," in *Proc. Int. Symp. Computer Vision*, Nov. 1995, pp. 115–120.
[47] K. Lange, *Optimization*. New York: Springer Verlag, 2004.
[48] J. Darbon, "Composants logiciels et algorithmes de minimisation exacte d'énergies dédiés au traitement des images," Ph.D. dissertation, Ecole Nationale Sup. Télécom., France, 2005.
[49] V. Kolmogorov, Primal-dual algorithm for convex markov random fields Cambridge, U.K., Tech. Rep., Microsoft Res., 2005.
[50] K. Murota, "On steepest descent algorithms for discrete convex functions," *SIAM J. Optim.*, vol. 14, no. 3, pp. 699–707, 2003.
[51] J. Darbon and M. Sigelle, "Image restoration with discrete constrained total variation part I: Fast and exact optimization," *J. Math. Imag. Vis.*, 2006.
[52] A. Chambolle, "Total variation minimization and a class of binary MRF models," in *Proc. 5th Int. Workshop*, St. Augustine, FL, Nov. 2005, vol. 3757, pp. 136–152.
[53] D. Hochbaum, "An efficient algorithm for image segmentation, Markov random fields and related problems," *J. ACM*, vol. 48, no. 2, pp. 686–701, 2001.
[54] D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posteriory estimation for binary images," *J. Roy. Statist. Soc. B*, vol. 51, no. 2, pp. 271–279, 1989.
[55] P. Ferrari, M. Gubitoso, and E. Neves, "Reconstruction of gray-scale images," *Method. Comput. Appl. Probab.*, vol. 3, pp. 255–270, 2001.
[56] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1101–1113, Nov. 1993.
[57] H. Ishikawa, "Exact optimization for Markov random fields with convex priors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1333–1336, Oct. 2003.
[58] D. Bertsekas, *Network Optimization: Continuous and Discrete Models*. Belmont, MA: Athena-Scientific, 1998.
[59] A. Goldberg and R. Tarjan, "A new approach to the maximum-flow problem," *J. Assoc. Comput. Mach.*, vol. 35, no. 4, pp. 921–940, Oct. 1988.
[60] S. Li, *Markov Random Field Modeling in Computer Vision*. New York, NJ: Springer-Verlag, 1995.
[61] K. Murota, *Discrete Convex Analysis*. Philadelphia, PA: SIAM, 2003.
[62] S. Geman and D. McClure, "Statistical methods for tomographic image reconstruction," in *Proc. 46th Session of the International Statistical Institute*, 1987, vol. 52, pp. 353–356.
[63] G. Nico, "Phase reconstruction via simulated tempering," *Phys. Lett. A*, vol. 280, pp. 386–395, 2001.

José M. Bioucas-Dias (SM'95) received the E.E., M.Sc., and Ph.D. degrees in electrical and computer engineering, all from the Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal, in 1985, 1991, and 1995, respectively.

He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, IST. He is also a researcher with the Communication Theory and Pattern Recognition Group of the Institute of Telecommunications. His research interests include remote sensing, signal and image processing, pattern recognition, and communications.

Gonçalo Valadão received the degree in engineering physics and the M.Sc. degree in geographical information systems (GIS) from the Instituto Superior Técnico (IST), Technical University of Lisbon, Lisbon, Portugal, in 2000, where he is currently pursuing the Ph.D. degree in electrical and computer engineering.

He was with a management/IT consultancy, after which he joined an R&D group at IST where he developed GIS. His research work is now being done within the Communication Theory and Pattern Recognition Group, Instituto de Telecomunicações, Lisbon, under the supervision of Prof. Bioucas-Dias. His research interests include phase unwrapping, and more broadly, remote sensing, image processing, and machine learning.