Joint Regularization of Phase and Amplitude of InSAR Data: Application to 3D reconstruction

Loïc Denis, Florence Tupin, Jérôme Darbon and Marc Sigelle

Abstract

Interferometric SAR images suffer from a strong noise and their regularization is often a prerequisite for successful use of their information. Independently of the unwrapping problem, interferometric phase denoising is a difficult task due to shadows and discontinuities. In this paper, we propose to jointly filter phase and amplitude data in a Markovian framework. The regularization term is expressed by the minimization of the total variation and may combine different information (phase, amplitude, optical data). First, a fast and approximate optimization algorithm for vectorial data is briefly presented. Then two applications are described. The first one is a direct application of this algorithm for 3D reconstruction in urban areas with Very High Resolution (VHR) images. The second one is an adaptation of this framework to the fusion of SAR and optical data. Results on aerial SAR images are presented.

I. Introduction

The two previous years have seen a new generation of SAR sensors (TerraSAR-X [20], ALOS, CSK [16], RadarSat-2) with increased resolution and smaller revisit time thanks to constellation. Although very popular for their all-weather and all-time capabilities and their polarimetric and interferometric potential, SAR data remain difficult to use and processing tools are still necessary to fully profit from them. In this paper, we propose a new filtering tool for interferometric SAR data and investigate its application for 3D reconstruction with and without an optical image.

This paper is divided in 3 parts. In the first section, a recent fast and approximate algorithm [11] for the optimization of Markov Random Fields (MRF) with convex prior and defined on vectorial data is recalled for the paper to be self consistent. Then, two applications of this algorithm are described which constitute the main contributions of the paper. The first one is the joint regularization of phase and amplitude data and their use for 3D building reconstruction. The method and associated results are described in section III. The second application is the adaptation of this framework to the fusion of optical and interferometric data. Methodology and results on real data are presented in section IV.

This work was supported by the Centre National d'Études Spatiales under the project R-S06/OT04-010.

L. Denis was with GET/Télécom Paris, CNRS UMR 5141, Paris, France, and is now with École Supérieure de Chimie Physique Électronique de Lyon and Laboratoire Hubert Curien, CNRS UMR 5516, St-Étienne.

F. Tupin and M. Sigelle are with GET/Télécom Paris, CNRS UMR 5141, Paris, France.

J. Darbon is with the Department of Mathematics of UCLA, Los Angeles, USA. Research supported by ONR Grant N000140710810 and also partially through NSF DMS-0610079 and ONR N00014-06-1-0345.

II. Fast and approximate algorithm for the optimization of TV with vectorial data

In this section, we briefly recall the optimization algorithm that will be used in the two following parts. A more complete description of the method together with a comparison with other optimization approaches may be found in [11].

A. MRF framework

It is assumed that a vectorial image $\mathbf{u}$ is defined on a finite discrete lattice S and takes values in a discrete multi-dimensional integer set $\mathcal{L} = \{1, \ldots, L\}^N$ (where N is the dimension, i.e., the number of channels per pixel). We denote by $\mathbf{u}_s$ the vectorial value of the image $\mathbf{u}$ at the site $s \in S$.

Given an observed image $\mathbf{u}$, a Bayesian analysis using the MAP criterion (Maximum A Posteriori) consists of finding a restored image $\hat{\mathbf{u}}$ that maximizes:

$$p(\hat{\mathbf{u}}|\mathbf{u}) \propto p(\mathbf{u}|\hat{\mathbf{u}})p(\hat{\mathbf{u}}).$$

It can be shown under the assumption of Markovianity of $\hat{\mathbf{u}}$ (with order-2 clique) and with some independence assumption on $\mathbf{u}$ conditionally to $\hat{\mathbf{u}}$ ($P(\mathbf{u}|\hat{\mathbf{u}}) = \Pi_s P(\mathbf{u}_s|\hat{\mathbf{u}}_s)$) that the MAP problem is an energy minimization problem:

$$\hat{\mathbf{u}}^{(MAP)} = \arg\min_{\hat{\mathbf{u}}} E(\hat{\mathbf{u}}|\mathbf{u}) \ ,$$

with

$$E(\hat{\mathbf{u}}|\mathbf{u}) = \sum_s U(\mathbf{u}_s|\hat{\mathbf{u}}_s) + \beta \sum_{(s,t)} \psi(\hat{\mathbf{u}}_s, \hat{\mathbf{u}}_t) \ ,$$

$U(\mathbf{u}_s|\hat{\mathbf{u}}_s) = -\log p(\mathbf{u}_s|\hat{\mathbf{u}}_s)$ and ψ is a function modeling the prior chosen for the solution. In this paper, the ψ functions considered will be convex functions of $\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t$, as required to apply the combinatorial optimization algorithm described below.

B. Energy minimization problem

Graph-cut based approaches are very efficient methods for MRF optimization. In some cases, they lead to the global minimum of non-convex problems with potentially numerous local minima. Exact optimization can be achieved by two methods. The first one has been developed by Ishikawa [13]. It is able to handle any kind of data driven term ($U(\mathbf{u}_s|\hat{\mathbf{u}}_s)$) provided the regularization term (ψ function) is convex. In the case of remote sensing however, the size of the graph to be built and stored in memory is prohibitive. The second method has been proposed in [8]. It is based on the notion of levelable energies, but again the memory size is prohibitive for remote sensing data. As far as approximate optimization methods are concerned, for regularization functions which are metric, α-expansion algorithm proposed in [3] can be applied. Starting from a current solution, this algorithm proposes to each pixel either to keep its current value, or to take α as new value. The energy associated to each of the possible changes is minimized through a minimum cut computation. The succession of α-expansions over all possible values in $\mathcal{L}$ until convergence leads to a solution which is shown to be close to the global minimum. Besides, this approach has been shown in [6] to converge to a global minimizer when data fidelity is convex. If the set of all possible pixel values can be large in the case of single image

regularization, its size becomes prohibitive when vectorial data are considered. We suggest in the next section a faster algorithm which is more suitable when large or multi-dimensional images are considered.

Minimizing a non-convex energy is a difficult task as the algorithm may fall into a local minimum. Algorithms such as the Iterated Conditional Modes [1] require a "good" initialization and then perform local changes to reduce the energy. Graph-cut approaches provide a way to explore a combinatorial set of changes involving simultaneously all pixels. Following [3], we denote such changes *large moves*. Instead of allowing a pixel to either keep its previous value or change it to a given one (α-expansion), we suggest that a pixel could either remain unchanged or its value be increased (or decreased) by a fixed step. Such an approach has been firstly described independently in [2], [5], [14] and applied recently with unitary steps in [2]. We however use these large moves in a case of non-convex data term. The trial steps are chosen to perform a scaling sampling of the set of possible pixel values. We express the algorithm in the general case of vectorial data.

We first describe the set of large moves considered, then the associated graph construction and the resulting algorithm.

1) Local minimization: First, let us introduce the set of images that lie within a single move in our algorithm.

$$\mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)}) = \{\hat{\mathbf{u}} \,/\, \forall s \in S, \exists k_s \in \{0,1\}, \hat{\mathbf{u}}_s = \hat{\mathbf{u}}_s^{(n)} + k_s \mathbf{d}\}$$

is the set of images whose pixel value $\hat{\mathbf{u}}_s$ is either unchanged or increased by step $\mathbf{d}$. We define the "best" move $\hat{\mathbf{u}}^{(n)} \mapsto \hat{\mathbf{u}}^{(n+1)}$ as the one that minimizes the restriction of the energy to the set $\mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)})$:

$$\hat{\mathbf{u}}^{(n+1)} = \underset{\hat{\mathbf{u}}^{(n+1)} \in \mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)})}{\arg\min} E(\hat{\mathbf{u}}^{(n+1)}|\mathbf{u}).$$

The restriction of the energy to $\mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)})$ corresponds to an energy involving only the binary variables $(k_s)_{s \in S}$. According to [15], an energy of binary variables arising from a first-order Markov model can be minimized by computing a minimum cut on a related graph provided it satisfies the following submodular property:

$$\psi(0,1) + \psi(1,0) \geq \psi(0,0) + \psi(1,1).$$

To compute the "best" move using a s-t minimum-cut algorithm, the following must therefore hold:

$$\psi(\hat{\mathbf{u}}_s, \hat{\mathbf{u}}_t + \mathbf{d}) + \psi(\hat{\mathbf{u}}_s + \mathbf{d}, \hat{\mathbf{u}}_t) \geq \psi(\hat{\mathbf{u}}_s, \hat{\mathbf{u}}_t) + \psi(\hat{\mathbf{u}}_s + \mathbf{d}, \hat{\mathbf{u}}_t + \mathbf{d}). \tag{1}$$

Note that in most cases, the prior model ψ depends only on the difference $\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t$. For such prior models, condition 1 becomes:

$$\psi(\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t - \mathbf{d}) + \psi(\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t + \mathbf{d}) \geq 2\psi(\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t)$$

which is the definition of the convexity of ψ.

In conclusion, the *local* problem of finding the vectorial field $\hat{\mathbf{u}}^{(n+1)}$ located within a single move (i.e., $\hat{\mathbf{u}}^{(n+1)} \in \mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)})$) that minimizes the posterior energy $E(\hat{\mathbf{u}}^{(n+1)}|\mathbf{u})$ can be *exactly* solved by computing a minimum cut on a graph (described in next paragraph) provided that the regularization potential is convex and depends only on the difference $\hat{\mathbf{u}}_s - \hat{\mathbf{u}}_t$.

2) Graph construction: We build a graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, following the method of [15], to minimize the restriction of the energy to allowed moves of step $\mathbf{d}$:

$$\underset{(k_s)_{s\in S}}{\arg\min} \sum_s U(\mathbf{u}_s|\hat{\mathbf{u}}_s^{(n)} + k_s\mathbf{d}) + \beta \sum_{(s,t)} \psi(\hat{\mathbf{u}}_s^{(n)} + k_s\mathbf{d}, \hat{\mathbf{u}}_t^{(n)} + k_t\mathbf{d}) \tag{2}$$

The graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is directed, with nonnegative edge weights and two terminal vertices: the source $\mathscr{S}$ and the sink $\mathscr{P}$. The graph structure and the edge weights are chosen such that any cut[1] has a cost (i.e., sum of the cut edges capacities) corresponding to the energy to minimize. We create a vertice for each site s, all connected respectively to the source and the sink through two edges with capacity $c_{s,1}$ (resp. $c_{s,0}$). Finally, each clique (s,t) gives rise to an edge with capacity $c_{s,t}$ (fig. 1).

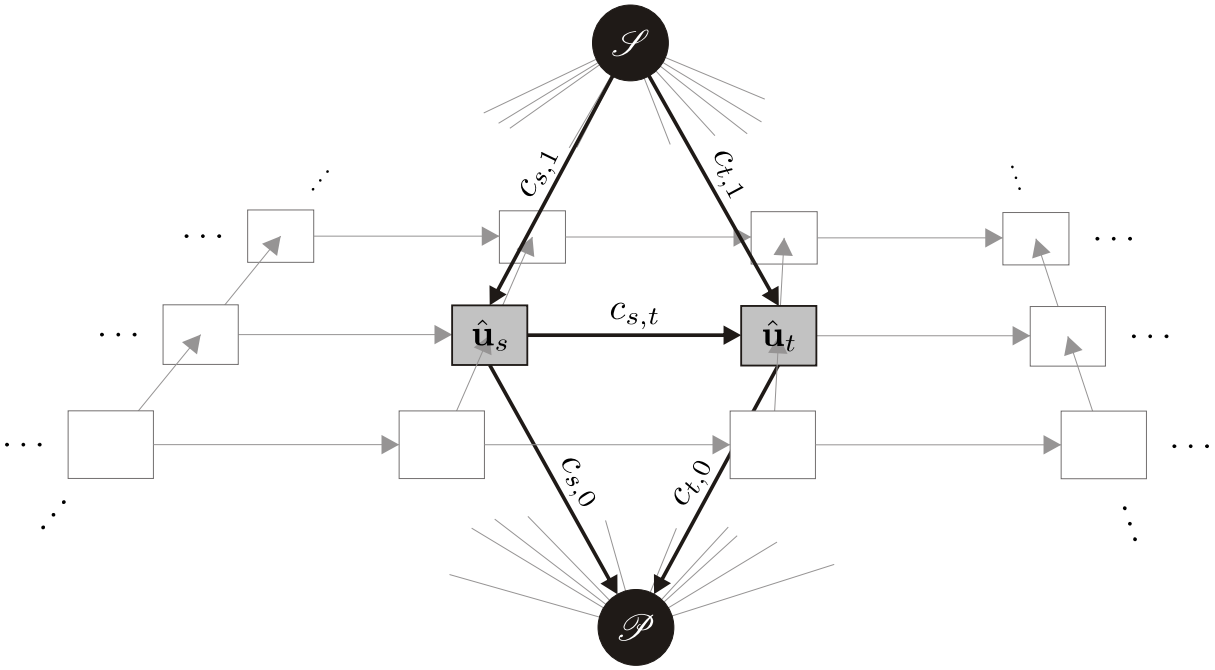


Fig. 1. Graph construction for local minimization: the graph has one layer of nodes (one per pixel) and two terminals (the source $\mathcal{S}$ and the sink $\mathcal{P}$).

The capacities are set according to the additive method described in [15]. The first term in equation (2) is represented by the weights:

$$\begin{cases} c_{s,1} = \max(0, U(\mathbf{u}_s|\hat{\mathbf{u}}_s^{(n)} + \mathbf{d})) - U(\mathbf{u}_s|\hat{\mathbf{u}}_s^{(n)})) \\ c_{s,0} = \max(0, U(\mathbf{u}_s|\hat{\mathbf{u}}_s^{(n)}) - U(\mathbf{u}_s|\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}))). \end{cases}$$

[1]a cut is a partition of the vertices into two disjoint sets S and P such that $\mathscr{S} \in \mathcal{S}$ and $\mathscr{P} \in \mathcal{P}$

We add to these weights other weights that correspond to the clique term (second term of equation 2):

$$\begin{cases} c'_{s,1} &= \beta \cdot \max\left(0, \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)}) - \psi(\hat{\mathbf{u}}_s^{(n)}, \hat{\mathbf{u}}_t^{(n)})\right) \\ c'_{s,0} &= \beta \cdot \max\left(0, \psi(\hat{\mathbf{u}}_s^{(n)}, \hat{\mathbf{u}}_t^{(n)}) - \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)})\right) \\ c'_{t,1} &= \beta \cdot \max(0, \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)} + \mathbf{d}) \\ & \quad -\psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)})) \\ c'_{t,0} &= \beta \cdot \max(0, \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)}) \\ & \quad -\psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)} + \mathbf{d})) \\ c_{s,t} &= \beta \cdot (\psi(\hat{\mathbf{u}}_s^{(n)}, \hat{\mathbf{u}}_t^{(n)} + \mathbf{d}) + \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)}) \\ & \quad -\psi(\hat{\mathbf{u}}_s^{(n)}, \hat{\mathbf{u}}_t^{(n)}) - \psi(\hat{\mathbf{u}}_s^{(n)} + \mathbf{d}, \hat{\mathbf{u}}_t^{(n)} + \mathbf{d})) \end{cases}$$

3) Approximate global minimization: When non-convex data terms are considered, the global minimization problem can not be exactly solved without considering each possible configuration (i.e., building a huge graph). On the other hand, when all terms are convex, it has been proven in [2], [5] that a succession of local minimizations leads to the global minimum. An exploration based on different scalings of the step is then suggested to speed up convergence.

We follow here a heuristic method that combines the *exact* determination of the best moves, with no guarantee on how close to the global minimum we get. In one dimension, a scaling search is performed by looking for the best move with steps $d_i^+ = L/2^i$ and $d_i^- = -L/2^i$ for i from 1 to the desired precision (i.e., quantization level). In N dimensions, there are $3^N - 1$ vectorial steps $\mathbf{d}_i$ to consider for a given step size d_i:

$$\mathbf{d}_i \in \mathscr{S}(d_i) \stackrel{\text{def}}{=} \{0, -d_i, +d_i\}^N / \{0, \ldots, 0\}.$$

The regularization algorithm for vectorial data is summarized here:

1: **for all** $s \in S$ **do**
2: $\quad \hat{\mathbf{u}}_s^{(0)} \leftarrow \{L/2, \ldots, L/2\}$
3: **end for**
4: $n \leftarrow 0$
5: **for** $i = 1$ to $precision$ **do**
6: $\quad d_i \leftarrow L/2^i$
7: $\quad$ **for all** $\mathbf{d}_i \in \mathscr{S}(d_i)$ **do**
8: $\quad\quad \hat{\mathbf{u}}^{(n+1)} \leftarrow \arg\min_{\hat{\mathbf{u}}^{(n+1)} \in \mathscr{S}_{\mathbf{d}}(\hat{\mathbf{u}}^{(n)})} E(\hat{\mathbf{u}}^{(n+1)} | \mathbf{u})$
9: $\quad\quad n \leftarrow n + 1$
10: $\quad$ **end for**
11: **end for**

Line 8 represents the exact binary energy minimization obtained by computing a minimum cut on a graph built according to section II-B2. Note that if we perform unitary steps $\mathbf{d}_i \in \mathscr{S}(1)$ until convergence at the termination of our algorithm, exact minimization is then guaranteed for convex energies [2], [5]. Results on the performances (speed, complexity, quality of the optimum) of the algorithm may be found in [11].

III. Application to 3D reconstruction with InSAR data

In this part, we will present an application of the previous algorithm to the joint filtering of phase and amplitude data for 3D reconstruction purpose [10]. We are interested in this paper in high resolution interferometric data. In many cases of interest (such as urban area imaging), the elevation range is contained within one fringe. The scene may then contain sharp transitions that must be well preserved (i.e., well modeled) in the reconstruction algorithm. We focus in this paper on such non-wrapped phase data.

A. Energy definition

1) Log-likelihood term: The synthesized radar image z is complex-valued. Its amplitude $|z|$ is very noisy due to the interferences that occur inside a resolution cell. Under the classical model of Goodman [12], the amplitude a_s of a pixel s follows a Nakagami distribution depending on the square root of the reflectivity $\hat{a}_s$. This likelihood leads to the following energetic term:

$$U(a_s|\hat{a}_s) = M\left[\frac{a_s^2}{\hat{a}_s^2} + 2\log\hat{a}_s\right]. \tag{3}$$

In the case of SAR interferometric data, the interferometric product is obtained by complex averaging of the hermitian product γ of the two SAR images. A good approximation of the phase ϕ_s distribution is a Gaussian which leads to a quadratic energy:

$$U(\phi_s|\hat{\phi}_s) = \frac{(\phi_s - \hat{\phi}_s)^2}{\hat{\sigma}^2_{\phi_s}}. \tag{4}$$

The standard deviation $\hat{\sigma}^2_{\phi_s}$ at site s is approximated by the Cramer-Rao bound $\hat{\sigma}^2_{\phi_s} = \frac{1-\rho_s^2}{2L\rho_s^2}$ (with L the number of average samples and ρ_s the coherence of site s). For low coherence areas (shadows or smooth surfaces, denoted $Shadows$ in the following), this Gaussian approximation is less relevant and a uniform distribution model is preferred: $p(\phi_s|\hat{\phi}_s) = \frac{1}{2\pi}$.

2) Regularization term: For mono-dimensional data, the minimization of the total variation is a very popular model [4], [7], [17]–[19] since the seminal work of [21]. It corresponds to a function ψ defined by the absolute value of the difference of neighboring pixels and has a behavior which preserves discontinuities in the regularized signal, while being convex. In the case of urban areas, many sharp discontinuities exist either in the amplitude image or in the interferometric one, so this model is well adapted.

The proposed method aims at preserving *simultaneously* phase and amplitude discontinuities. Indeed, the phase and amplitude information are hopefully linked since they reflect the same scene. Amplitude discontinuities are thus usually located at the same place as phase discontinuities and conversely. We propose in this paper to perform the joint regularization of phase and amplitude. To combine the discontinuities a disjunctive $\max$ operator is chosen. The joint prior model is defined by:

$$E(\hat{a}, \hat{\phi}) = \sum_{(s,t)} \max(|\hat{a}_s - \hat{a}_t|, \gamma|\hat{\phi}_s - \hat{\phi}_t|), \tag{5}$$

with γ a parameter that can be set to 1, and otherwise accounts for the relative importance given to the discontinuities of the phase ($\gamma > 1$) or of the amplitude ($\gamma < 1$). As requested by the minimization algorithm, this regularization potential is convex.

3) Global energy: The global joint energy term is then defined as:

$$E(\hat{a},\hat{\phi}|a,\phi) = \frac{1}{\beta_a}\sum_s M[\frac{a_s^2}{\hat{a}_s^2} + 2\log\hat{a}_s]$$
$$+\frac{\gamma}{\beta_\phi}\sum_s \frac{(\phi_s-\hat{\phi}_s)^2}{\hat{\sigma}^2_{\phi_s}} + \sum_{(s,t)} \max(|\hat{a}_s-\hat{a}_t|, \gamma|\hat{\phi}_s-\hat{\phi}_t|).$$

β_a and β_ϕ are some weightings of the likelihood terms introduced in order to balance the data fidelity and regularization terms.

4) Processing of shadow areas: Due to the specific properties of shadow areas (random phase implying no likelihood term), they are separately detected and an adapted regularization term is defined. The regularized fields $\hat{a}$ and $\hat{\phi}$ at sites s located inside the detected shadow areas $Shadows$ are governed only by the regularization term. With the prior term defined in equation (1), the phase $\hat{\phi}_s$ for $s \in Shadows$ that minimizes the energy corresponds to an interpolation of the phase value at the surrounding sites. However shadows areas are most of the time at ground level and not at an intermediate height between the top of the structure that creates the shadow and the ground at the shadow end. A modified regularization term that better describes this prior knowledge is therefore used for cliques involving one or both site(s) inside the shadow regions: $E(\hat{a},\hat{\phi}) = \sum_{(s,t)} E(\hat{a},\hat{\phi})_{(s,t)}$ with $E(\hat{a},\hat{\phi})_{(s,t)}$ defined as:

(i) if $s \notin Shadows$ and $t \notin Shadows$,
$E(\hat{a},\hat{\phi})_{(s,t)} = \max(|\hat{a}_s-\hat{a}_t|, \gamma|\hat{\phi}_s-\hat{\phi}_t|)$,

(ii) if $s \in Shadows$ and $t \notin Shadows$ and $\hat{\phi}_s \leq \hat{\phi}_t$
$E(\hat{a},\hat{\phi})_{(s,t)} = |\hat{a}_s-\hat{a}_t| + \gamma|\hat{\phi}_s-\hat{\phi}_t|$,

(iii) if $s \in Shadows$ and $t \notin Shadows$ and $\hat{\phi}_s > \hat{\phi}_t$
$E(\hat{a},\hat{\phi})_{(s,t)} = |\hat{a}_s-\hat{a}_t| + 2\gamma|\hat{\phi}_s-\hat{\phi}_t|$,

(iv) if $s \in Shadows$ and $t \in Shadows$
$E(\hat{a},\hat{\phi})_{(s,t)} = |\hat{a}_s-\hat{a}_t| + \gamma\left(\hat{\phi}_s-\hat{\phi}_t\right)^2$.

The cases where $s \notin Shadows$ and $t \in Shadows$ are treated in a symmetrical manner. Outside shadow areas (case i), the regularization term is the same as previously. To limit the effect of a given shadow area on the regularization of the amplitude, we independently regularize phase and amplitude in and at the limit of the shadows (cases ii to iv). To force the regularized phase inside a shadow to follow ground level, we penalize more heavily over-estimation (case iii) than under-estimation (case ii). Finally, a quadratic constraint (case iv) enforces a smooth ground inside a shadow area. Note that in each case (i to iv) the prior term $E(\hat{a},\hat{\phi})_{(s,t)}$ is convex and so is the prior energy $E(\hat{a},\hat{\phi})$.

B. Results

The proposed joint regularization model and the fast and approximate regularization algorithm have been applied to two VHR interferometric images. A result is presented in figure 3 for the regularized phase and amplitude and in figure 2 for a 3D view.

The results are very close to the ones obtained in [23] with a very different approach. Note that a weaker hypothesis is made here, since in [23] an assumption of planar surface is done for each region on the optical

Fig. 2. 3D view of the amplitude superimposed on the elevation derived from the regularized phase.

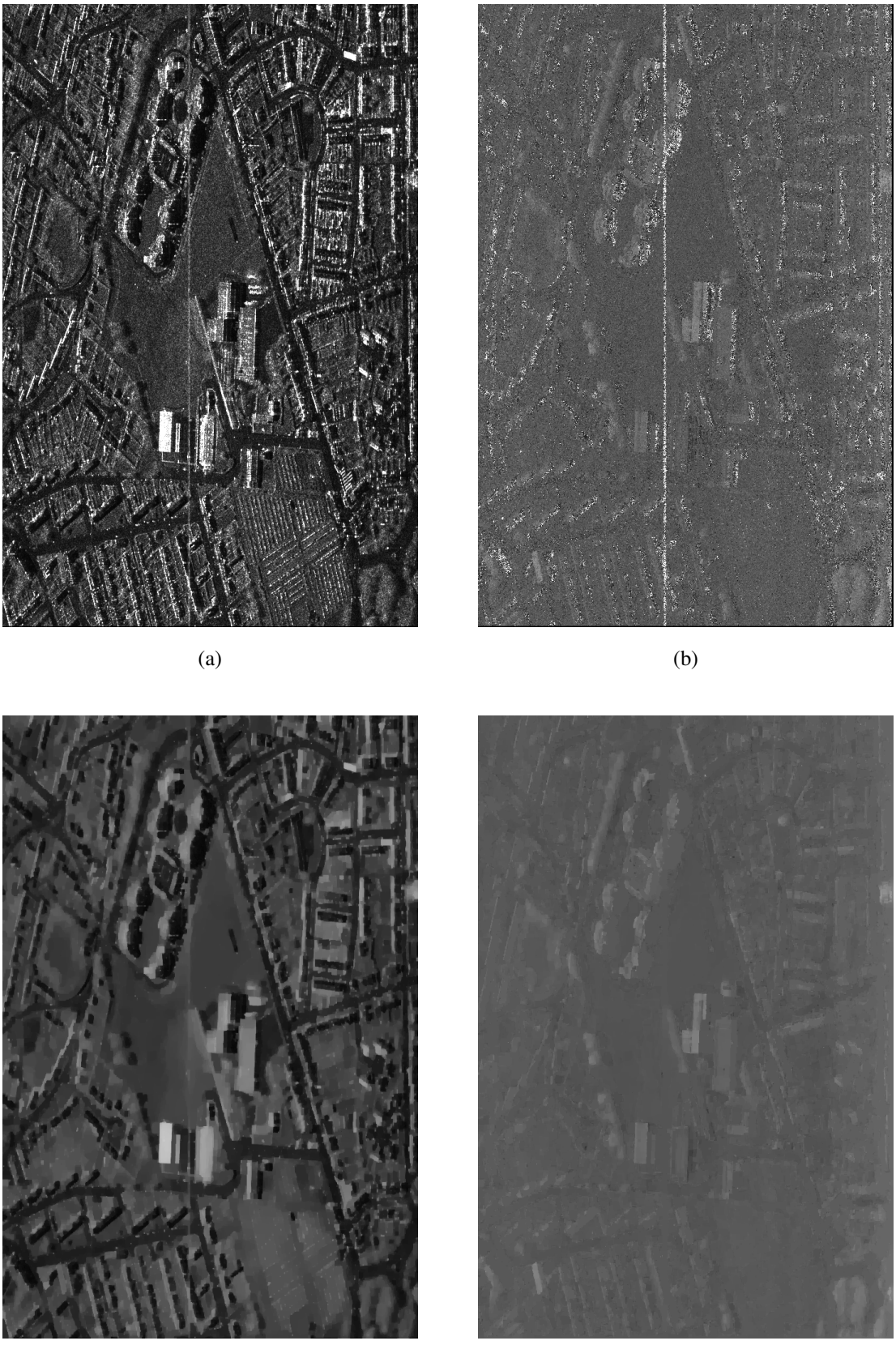
image. However, a precise comparison is difficult due to the sensor parameters. Indeed, the baseline is 0.7m leading to an ambiguity altitude of 180m and an altimetric precision between 2 and 3 meters. The accuracy of the height retrieval for the large buildings is satisfying (a root mean square error of 2.5 m has been computed) compared to the given altimetric precision.

IV. Application to 3D reconstruction with InSAR and optical data fusion

The second application is dedicated to the fusion of InSAR and optical data [9]. One of the main problems is the projection of the SAR pixels using the elevation provided by the phase image in the optical geometry. After projection, the cloud of points is irregular and our approach relies on a triangulation of this cloud.

A. Description of the method

The suggested method is summarized in figure 4. The main steps, denoted with circled numbers on the figure, are the following:



(a) (b)

(c) (d)

Fig. 3. Original SAR image (on the left amplitude and phase on the right) ©DGA/CNES and their joint regularization (below).

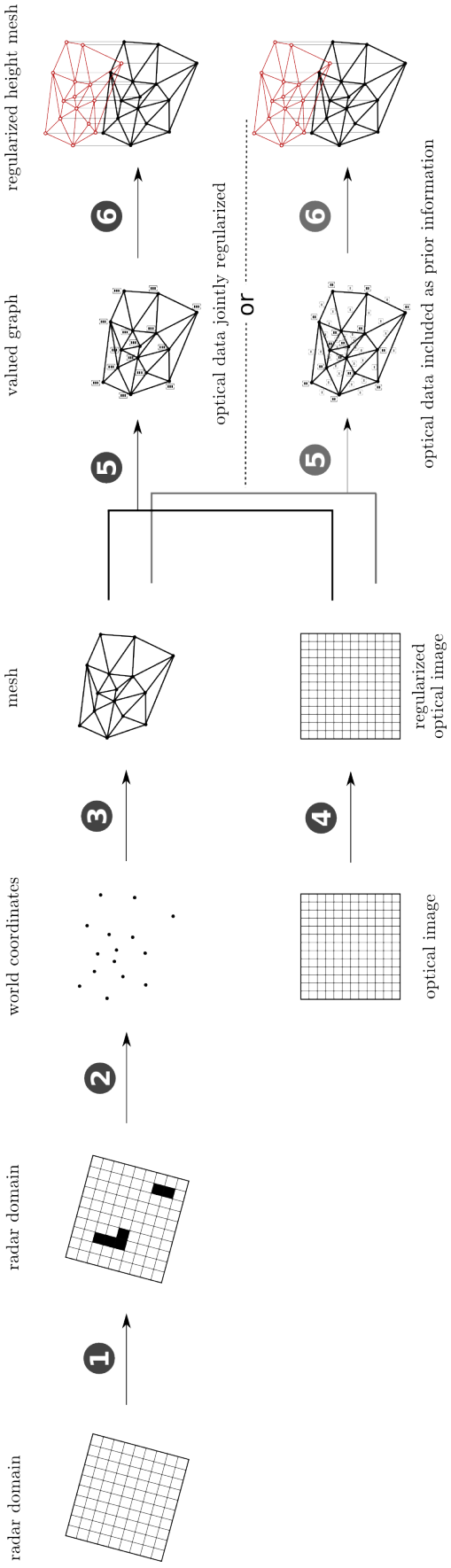


Fig. 4. Scheme of the suggested method. The numbers correspond to the algorithm steps referred to in the text.

❶ The shadows are detected on the radar image.

❷ The height map in the world coordinates is obtained by projection of all points from the radar image except those in shadow areas.

❸ The cloud of points of the height map is then triangulated into a mesh. This mesh defines the topology of the graph used in steps ❺ and ❻.

❹ To ease the introduction of optical information, the optical image is regularized prior to graph construction.

❺ A valued graph is then built with nodes corresponding to each of the points in the cloud, arcs given by the mesh and values set using the SAR amplitude, height and the optical information.

❻ Once the graph is built, a regularized height mesh is computed by defining a Markov field over the graph.

B. Preprocessing steps

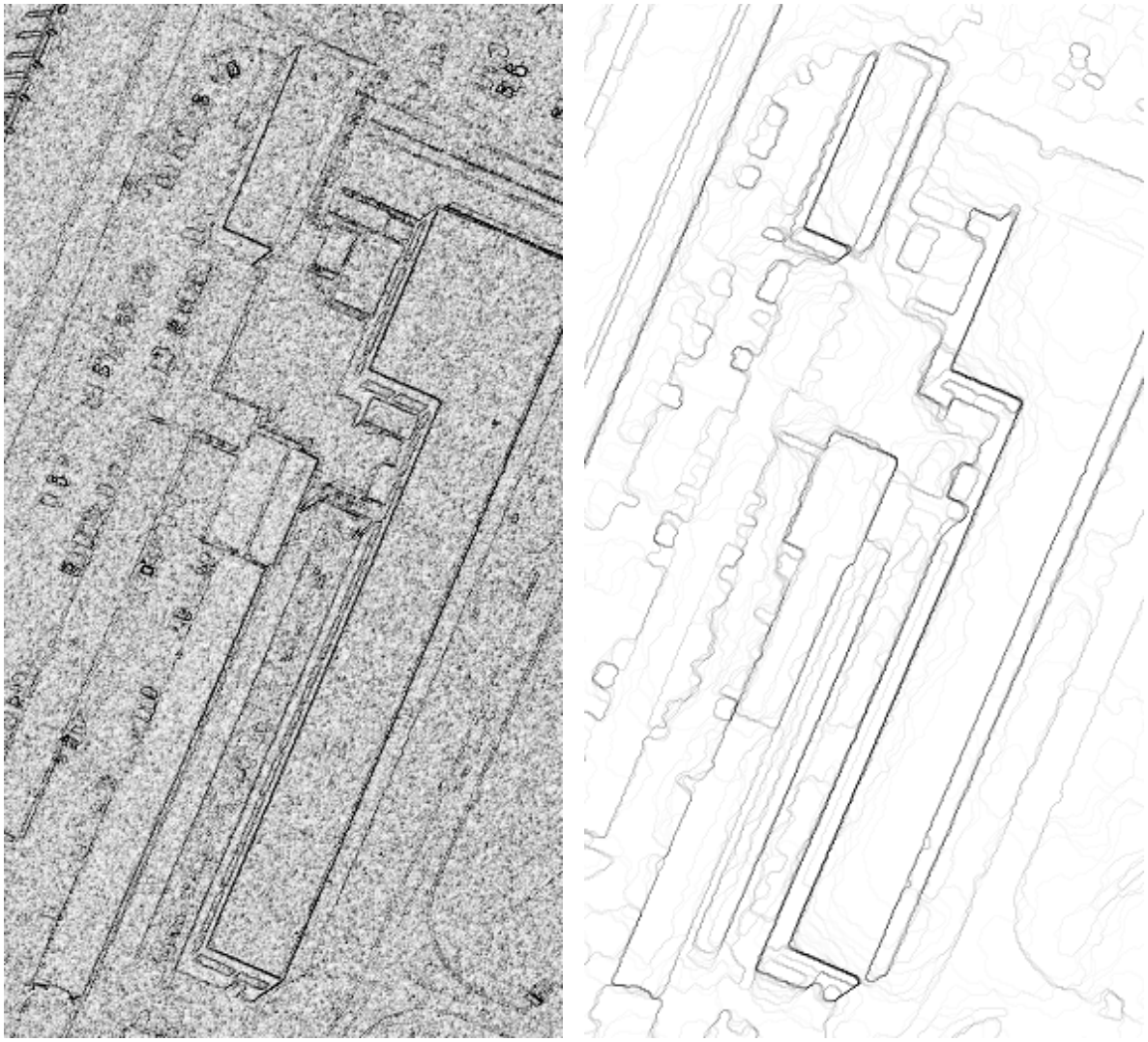
Steps ❶ to ❹ are preprocessing steps required before the actual height regularization (steps ❺–❻). Before merging the InSAR and optical data to perform a 3D reconstruction, images must be transformed into a common coordinate system. Assuming the optical image is acquired at normal incidence, we then have to project back the InSAR data from distance sampling coordinates to 3D coordinates. Before projecting the points from radar geometry to world coordinates, shadows are detected (step ❶) to prevent from projecting points with unknown (i.e., random) height. This detection is made using the Markovian classification described in [22]. Points outside the shadows are then projected based on their interferometric phase and the radar acquisition parameters (step ❷). This gives a 3D cloud of points (x, y, z) in the world coordinates. The projection of this cloud on a horizontal plane is then triangulated with Delaunay algorithm to obtain a height mesh (step ❸). The height of each node of the obtained graph can then be regularized (see next section). The optical image is simplified using a geometry+texture decomposition [24] before fusion (step ❹). This decomposition is obtained with a TV+L1 regularization computed using the graph cut algorithm described in section II. Figure 5 displays the gradient norm of the optical image before and after its regularization. Most irrelevant edges are removed.

C. Height regularization model

In this application, the joint information of amplitude and interferometric data is used together with the optical data. Similarly as in the previous section, we define the regularized height field as that which maximizes the posterior probability according to the log-likelihood and prior models described below.

1) Log-likelihood model: The model is exactly the same as the one used in the previous section for amplitude and phase log-likelihood (see equations 3 and 4).

2) Prior model: We devise a prior model that accounts for the phase and amplitude dependency and that introduces the edges of the optical image. Two options are possible to regularize the height mesh. The first option is to jointly regularize the phase, amplitude and optical images (steps ❺–❻ shown at the top row of figure 4) by including the optical data in the log-likelihood term and extending equation 5 to include the regularized optical image (i.e., use a ternary $\max$ operator). This solution requires to set adequately the weights of each of the terms. The second option consists of introducing the optical image gradient as a prior (in this



(a) (b)

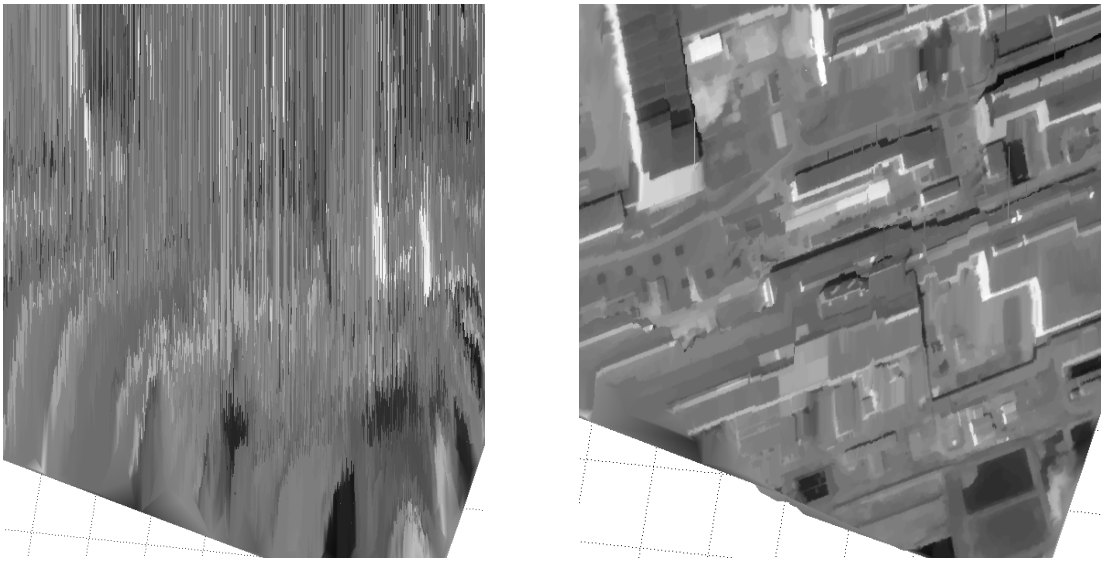
Fig. 5. Optical image regularization with TV+L1 decomposition model: (a) gradient norm of the optical image before regularization; (b) after regularization, remaining gradients correspond to the building edges.

case the optical image can be seen as an external field). Equation 5 is then replaced by:

$$E(\hat{a},\hat{\phi}) = \sum_{(s,t)} G_{opt}(s,t)\max(|\hat{a}_s - \hat{a}_t|, \gamma|\hat{\phi}_s - \hat{\phi}_t|) \tag{6}$$

with $G_{opt}(s,t) = \max(0, 1 - k_{opt}|opt_s - opt_t|)$ an expression that depends on the difference between the values opt_s and opt_t of the optical image at sites s and t, and k_{opt} a thresholding parameter.

When the optical image is constant between sites s and t, the $G_{opt}(s,t)$ term equals 1 and does not modify the joint TV regularization. When $|opt_s - opt_t|$ is high (corresponding to a discontinuity), $G_{opt}(s,t)$ is low, thus reducing the regularization of amplitude and phase. This modification permits to preserve the building shapes according to the optical data. Again, since the regularization potential is convex, the vectorial algorithm of section II can be applied.



(a) (b)

Fig. 6. Height mesh obtained (a) by direct projection of the InSAR data points (end of step ❸, fig. 4); (b) after joint phase/amplitude regularization with a prior that includes the optical image gradient (end of step ❻, second row of fig. 4).

D. Results and discussion

Figure 6(a) shows a height mesh with the regularized optical image used as texture. The mesh is too noisy to be usable. We performed a joint amplitude/phase regularization using the gradient of the optical image as a weight that eases the apparition of edges at the location of the optical image contours. The obtained mesh is displayed on figure 6(b). The surface is a lot smoother with sharp transitions located at the optical image edges. Buildings are clearly above the ground level (note that the shadows of the optical image create a fake 3D impression).

This approach requires a very good registration of the SAR and optical data and implies the knowledge of all acquisition parameters which is not always possible (depending on the source of images). The optical image should be taken with normal incidence to match the radar data. The image displayed on figure 6 was taken with a slight angle that displaces the edges and/or duplicates them. For the method to perform well, the edges of structures must be visible in both optical and InSAR images. A more robust approach would require a higher level analysis (significant edge detection, building detection).

V. Conclusion

In this paper we have presented two applications of a new optimization algorithm adapted for vectorial data. The first one is dedicated to the joint filtering of phase and amplitude for 3D reconstruction. The second one is the extension of such an approach for the introduction of information derived from the optical image. The framework described is quite general and can be used to fuse heterogeneous data according to their statistical distribution and to prior knowledge that can be introduced by various ways (edge co-location by joint regularization, variable weights, ...). The discrete minimization algorithm can handle energies with non-convex data-fidelity terms and (possibly non-smooth) convex priors. Such energies arise when modelling speckle noise

(non-convex log likelihood) and edge-preserving regularization using L1 norm. By defining the regularized field over a graph, it is possible to merge images with different sampling/geometry. In both cases promising results are obtained. Further work will be dedicated to a more extensive testing and evaluation of this framework, specially for the new data that are acquired by TerraSAR-X and CosmoSkyMed. Particularly, the phase filtering should be adapted to take wrapped phase into account. Concerning the fusion application, one of the main point to be investigated is the accurate registration between optical and SAR sensors with metric resolution.

REFERENCES

[1] J. Besag. On the statistical analysis of dirty pictures. *J. R. Statist. Soc. B*, 48(3):259–302, 1986.

[2] J. M. Bioucas-Dias and G. Valadão. Phase unwrapping via graph cuts. *IEEE Transactions on Image Processing*, 16(3):698–709, 2007.

[3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2001.

[4] P.L. Combettes and J.C. Pesquet. Image restoration subject to a total variation constraint. *IEEE Transactions on Image Processing*, 13(9):1213–1222, 2004.

[5] J. Darbon. *Composants logiciels et algorithmes de minimisation exacte d'énergies dédiées au traitement des images*. PhD thesis, Ecole Nationale Supérieure des Télécommunications (ENST E050), 2005.

[6] J. Darbon and S. Peyronnet. A vectorial self-dual morphological filter based on total variation minimization. In *Proceedings of the First International Conference on Visual Computing*, volume 3804 of *Lecture Notes in Computer Science Series*, pages 388–395, Lake Tahoe, Nevada, USA, December 2005. Springer-Verlag.

[7] J. Darbon and M. Sigelle. Image restoration with discrete constrained Total Variation part I: Fast and exact optimization. *Journal of Mathematical Imaging and Vision*, 26(3):261–276, December 2006.

[8] J. Darbon and M. Sigelle. Image restoration with discrete constrained Total Variation part II: Levelable functions, convex priors and non-convex cases. *Journal of Mathematical Imaging and Vision*, 26(3):277–291, December 2006.

[9] L. Denis, F. Tupin, J. Darbon, and M. Sigelle. A regularization approach for InSAR and optical data fusion. In *IGARSS'08*, Boston, USA, jul 2008.

[10] L. Denis, F. Tupin, J. Darbon, and M. Sigelle. Joint filtering of SAR interferometric phase and amplitude data in urban areas by TV minimization. In *IGARSS'08*, Boston, USA, jul 2008.

[11] L. Denis, F. Tupin, J. Darbon, and M. Sigelle. SAR image regularization with fast approximate discrete minimization. *IEEE Transactions on Image Processing (in revision)*, 2008.

[12] J.W Goodman. Statistical properties of laser speckle patterns. In *Laser Speckle and Related Phenomena*, volume 9, pages 9–75. J.C Dainty (Springer Verlag, Heidelberg, 1975), 1975.

[13] H. Ishikawa. Exact optimization for Markov random fields with convex priors. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(10):1333–1336, October 2003.

[14] V. Kolmogorov. Primal-dual algorithm for convex markov random fields. Technical report, Microsoft Research, 2005.

[15] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph-cuts ? *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(2), 2004.

[16] F. Lombardini, F. Bordoni, and F. Gini. Feasibility study of along-track SAR interferometry with the COSMO-Skymed satellite system. In *IGARSS04*, volume 5, pages 337–3340, 2004.

[17] Y. Meyer. *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*. American Mathematical Society, 2001.

[18] M. Nikolova. A variational approach to remove outliers and impulse no ise. *Journal of Mathematical Imaging and Vision*, 20:99–120, 2004.

[19] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An Iterative Regularization Method for Total Variation Based Image Restoration. *SIAM journal on Multiscale modeling and Application*, 4:460–489, 2005.

[20] R. Romeiser and H. Runge. Theoretical evaluation of several possible along-track InSAR modes of TerraSAR-X for ocean current measurements. *IEEE Transactions on Geoscience and Remote Sensing*, 45:21–35, 2007.

[21] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D.*, 60:259–268, 1992.

[22] C. Tison, J.M. Nicolas, F. Tupin, and H. Maître. A New Statistical Model of Urban Areas in High Resolution SAR Images for Markovian Segmentation. *IEEE Transactions on Geoscience and Remote Sensing (IGARSS'03 special issue)*, 42(10):2046–2057, oct 2004.

[23] C. Tison, F. Tupin, and H. Maître. A fusion scheme for joint retrieval of urban map and classification from high resolution interferometric SAR images. *IEEE Transactions on Geoscience and remote Sensing*, 45(2):495–505, mar 2007.

[24] W. Yin, D. Goldfarb, and S. Osher. A comparison of three total variation based texture extraction models. *Journal of Visual Communication and Image Representation*, 18(3):240–252, 2007.