THREE DIMENSIONAL RECONSTRUCTION OF URBAN AREAS USING JOINTLY PHASE AND AMPLITUDE MULTICHANNEL IMAGES

Aymen Shabou, Florence Tupin

Institut TELECOM, TELECOM ParisTech
CNRS-LTCI
Rue Barrault, 75013 Paris, France

Giampaolo Ferraioli, Vito Pascazio*

Università di Napoli Parthenope
Dipartimento per le Tecnologie
CdN, Isola C4, 80144 Napoli, Italy

1. INTRODUCTION

Three Dimensional (3D) reconstruction of urban areas is becoming a task of increasing importance in the last years, thanks to the high resolution of the new SAR sensors (TerraSAR-X, ALOS, CSK, RadarSat-2). The 3D reconstruction is performed using SAR Interferometry that exploits the phase of two or more SAR complex images.
Interferometric SAR images suffer from two main problems: they are often characterized by strong noise and they are wrapped in the main interval $[-\pi, \pi[$. Regularizing the phase images and solving the unwrapping problem are mandatory steps to provide the 3D reconstruction. Many works have been dedicated to the problem of interferometric phase filtering and to the unwrapping one. An effective way to combine these two problems is provided by the approach of [1]. In the latter, the filtering is directly incorporated into the unwrapping step, using a multichannel configuration which consists of exploiting more than two interferograms that are obtained using different frequencies or different baselines (i.e. different channels).
In this paper we are interested in the 3D reconstruction of urban areas using Very High Resolution (VHR) images. We use a multichannel configuration, as in [1], and we add to the 3D reconstruction chain the information provided by the amplitude data. We follow an approach similar to the one proposed in [2], where the authors jointly regularize the phase and the amplitude to provide an improved regularized phase. Differently from [2], in this paper we exploit the amplitude, not only to improve the phase regularization, but also to improve the phase unwrapping step.
We consider height profiles characterized by an elevation which is not contained within one fringe. The main idea consists of using the amplitude information to perform the 3D reconstruction, since in urban areas most height discontinuities (i.e., interferometric phase discontinuities) are also accompanied by amplitude discontinuities in SAR images and conversely (it is likely that the edges of one image are also present in the other one). The amplitude data provide useful information to preserve phase differences when height discontinuities appear and in the same time help to smooth noise in homogeneous areas.
To combine and exploit both amplitude and phase information we use a Markovian approach. The energy function is defined through a jointly likelihood statistical model of the amplitude and the interefromteric phase and a joint prior regularization function that allows us to preserve the edges and encourage their co-location in the restored amplitude and phase images. The optimization step is performed using a proposed Graph Cut based optimization algorithm.
In section 2, the proposed joint energy model combining both the likelihood and the prior terms is introduced. The algorithm used for the optimization step is addressed in section 3. Finally, some results on simulated data are shown in section 4.

2. ENERGY MODEL

Let us consider two complex SAR images $z_m^{(1)}$ and $z_m^{(2)}$, acquired at a certain frequency m. We will address the problem only in the multifrequency case, since the multibaseline one is almost equivalent. The interferometric phase of the two images can be computed as $\varphi_m = \angle(z_m^{(1)}\text{conj}(z_m^{(2)}))$, while the intensities images are given by $I_m^{(1)} = |z_m^{(1)}|^2$, $I_m^{(2)} = |z_m^{(2)}|^2$ and $I_m^{(1,2)} = |z_m^{(1)}\text{conj}(z_m^{(2)})|^2$. Given the intensity images and the interferometric phase for M different and independent channels, we want to jointly estimate the true amplitude image a and the true phase image ϕ using a Markovian approach. Let us consider the likelihood and the prior terms of the Markovian energy.

*Giampaolo Ferraioli performed the work while at TELECOM ParisTech, Départment TSI, Paris, France

2.1. Likelihood term

The joint negative multichannel log-likelihood of the intensity and interferometric phase at a given pixel p is a function of the unknown parameters a_p and ϕ_p. The expression is given by

$$E_p\left(I^{(1)}_{p,m}, I^{(2)}_{p,m}, I^{(1,2)}_{p,m}, \varphi_{p,m}|a_p, \phi_p\right) = \sum_{m=1}^{M}\left(4\log a_p + \frac{I^{(1)}_{p,m} + I^{(2)}_{p,m} - 2I^{(1,2)}_{p,m}\cdot\lambda_{p,m}\cdot\cos(\phi_p - \varphi_{p,m})}{a_p^2\cdot(1-\lambda_{p,m})}\right), \quad (1)$$

where $\lambda_{p,m}$ is the coherence value related to the p-th pixel of the m-th channel.

2.2. Prior term

To describe the prior term we use the Total Variation (TV) model. Since we are interested to urban areas that are characterized by many sharp transitions in both phase and amplitude images, the TV model results to be effective. As a matter of fact, the main feature of TV model is its effectiveness to preserve discontinuities without excessively penalizing smooth functions.
As stated before, amplitude discontinuities usually have the same location as phase discontinuities and conversely. Thus, we combine the discontinuities using a disjunctive max operator [2]. The joint prior model is defined by

$$E_{p,q}(a_p - a_q, \phi_p - \phi_q) = \max(\beta_a|a_p - a_q|, \beta_\phi|\phi_p - \phi_q|)\,, \quad (2)$$

where q is the index related to one of the 4 nearest pixel of the pixel p, while β_a and β_ϕ are two hyperparameters used to balance the amount of smoothing in the regularized phase and amplitude fields.
For the simplicity of notations, we write the total energy function as

$$E(\boldsymbol{x}|\boldsymbol{y}) = \sum_{p=1}^{N} E_p(\boldsymbol{x}_p) + \sum_{(p\sim q)} E_{p,q}(\boldsymbol{x}_p - \boldsymbol{x}_q)\,, \quad (3)$$

where $\boldsymbol{x} = [a^T\phi^T]^T$ is the unknown vector collecting both restored amplitude and profile values, $\boldsymbol{y} = [{I^{(1)}}^T {I^{(2)}}^T {I^{(1,2)}}^T \varphi^T]^T$ is the observed vector collecting the observed multichannel amplitude data and interferograms, N is the size of the restored data, $(p \sim q)$ denotes neighboring sites p and q, $E_p(\boldsymbol{x}_p) = \sum_{m=1}^{M} E_p(\boldsymbol{y}_{p,m}|\boldsymbol{x}_p)$ denoting the likelihood terms defined in (1) and $E_{p,q}(\boldsymbol{x}_p - \boldsymbol{x}_q) = E_{p,q}(a_p - a_q, \phi_p - \phi_q)$ denoting the regularization terms defined in (2). We will note for the rest of this paper by $\mathcal{L}_a$ and $\mathcal{L}_\phi$ respectively the discreete label sets of the reconstructed amplitude and profile images.
Solving the problem of joint multichannel phase and amplitude regularization consists of finding the optimal vectorial solution $\boldsymbol{x}$ that minimizes the non-convex multilabel energy function (3).

3. FAST APPROXIMATE MINIMIZATION

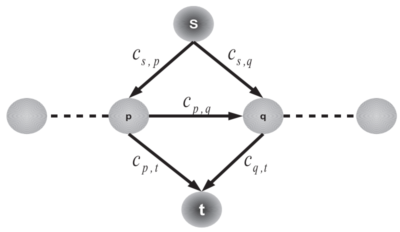
3.1. Vectorial expansion move

Minimizing non-convex multilabel energy function is a difficult task since the algorithm may fall into a local minimum. Classic algorithms like the Iterated Conditional Modes requires a good initialization and usually obtained results are not satisfying. Graph-cut based multilabel MRF optimization approaches, proposed firstly in [3], provide a way to explore effectively a combinatorial set of configurations while good local minima are provided. One of the most popular algorithms based on these approaches is the α-expansion algorithm. This approximate optimization algorithm is iterative and based on the concept of α-expansion move. The latter consists on changing a current configuration by proposing to any set of image pixels to change their labels to the label α. It finds a new configuration that minimizes the energy over all labelings within this move by building a particular graph and computing the s-t minimum cut/maximum flow on it. Expansion moves are iterated for all possible labels until convergence of the algorithm toward a local minimum. Computing maximum flow on the graph in a polynomial time needs submodularity condition on the regularization terms [3]. Thus, to be graph representable, the multilabel *a priori* function needs to be a metric.
As we are dealing with vectorial data, where the searched optimal configuration consists of two images, the amplitude and the phase ones, we propose in this work an extension of the α-expansion algorithm to this kind of vectorial data. We define a vectorial $\boldsymbol{\alpha}$-expansion move as the move from a configuration to another that minimizes the total energy (3) when a vector of labels $\boldsymbol{\alpha} = [\alpha_a\alpha_\phi]^T$ is proposed to current both amplitude and phase configurations. The minimization algorithm iterates for all possible couple of labels $(\alpha_a, \alpha_\phi) \in \mathcal{L}_a \times \mathcal{L}_\phi$, until convergence to a local minimum.

An optimal $\boldsymbol{\alpha}$-expansion move is performed based on the graph-cut technique. Thus, a specific directed graph is built with non-negative edge weights, where the s-t minimum cut leads to an optimal labeling of both amplitude and phase images. As we know, to perform the maximum flow computation, the submodularity of the *a priori* terms is needed, which is not the case of the proposed *a priori* function (2). A first possible solution that we propose to overcome this problem consists on restricting the move to submodular configurations. In this work, the performed expansion move proposes to all couple of pixels (a_p, ϕ_p) either to keep their current labels or to change *together* to the couple of labels (α_a, α_ϕ). In other words, the $\boldsymbol{\alpha}$-expansion move consists on proposing to the vector $\boldsymbol{x}$ either to keep it current value or to change to the vector $\boldsymbol{\alpha}$. Then, submodular configurations are considered.

3.2. Graph topography

In each single $\boldsymbol{\alpha}$-expansion move, a graph $\mathcal{G}_{\boldsymbol{\alpha}} = (\mathcal{V}_{\boldsymbol{\alpha}}, \mathcal{E}_{\boldsymbol{\alpha}})$ is created, where $\mathcal{V}_{\boldsymbol{\alpha}}$ is the set of vertices and $\mathcal{E}_{\boldsymbol{\alpha}}$ is the set of directed edges. We create a vertice for each site p. All vertices are connected to two special vertices: the source s and the sink t (for max-flow computation). Two families of directed edges connecting these vertices are defined: data edges that are related to the joint multichannel log-likelihood terms and interaction edges that are related to the joint *a priori* terms. We note with $c_{s,p}$ and $c_{p,t}$ the capacity of the data edge connecting the vertice p to the source and to the sink respectively and by $c_{p,q}$ the capacity of the interaction edge connecting neighboring pixels p and q. This graph structure and the edge weights are chosen such that any cut on $\mathcal{G}_{\boldsymbol{\alpha}}$ has a cost corresponding to the energy to minimize. Then, the s-t minimum cut gives the configuration with minimum energy. This configuration is obtained by assigning the label α_a (resp. α_ϕ) to the amplitude pixels p (resp. the phase pixels p) having the corresponding data edges connected to the source s in the cut, and for the rest of pixels, the current labeling of both amplitude and phase data remains unchanged.
In figure 1, a one dimensional $(1D)$ illustration of the graph $\mathcal{G}_{\boldsymbol{\alpha}}$ is presented where the construction for a neighboring pixels p and q is highlighted and capacity values for the created edges are given.



$$\begin{cases} c_{s,p} = max\{0, E_p(\boldsymbol{\alpha}) - E_p(\boldsymbol{x}_p)\} + max\{0, E_{p,q}(\boldsymbol{\alpha} - \boldsymbol{x}_q) - E_{p,q}(\boldsymbol{x}_p - \boldsymbol{x}_q)\}\,, \\ c_{p,t} = max\{0, E_p(\boldsymbol{x}_{\boldsymbol{p}}) - E_p(\boldsymbol{\alpha})\} + max\{0, E_{p,q}(\boldsymbol{x}_p - \boldsymbol{x}_q) - E_{p,q}(\boldsymbol{\alpha} - \boldsymbol{x}_q)\}\,, \\ c_{s,q} = E_{p,q}(\boldsymbol{\alpha} - \boldsymbol{x}_q)\,, \\ c_{q,t} = 0\,, \\ c_{p,q} = max\{0, E_{p,q}(\boldsymbol{\alpha} - \boldsymbol{x}_q) + E_{p,q}(\boldsymbol{x}_p - \boldsymbol{\alpha}) - E_{p,q}(\boldsymbol{x}_p - \boldsymbol{x}_q)\}\,. \end{cases}$$

(a) Graph structure. (b) Edge weights for a neighboring vertices p and q.

Fig. 1. A $1D$ illustration of the part of the graph on two neighboring vertices p and q. To obtain the total graph, we repeat this construction for all neighboring vertices, according to the neighborhood system.

4. RESULTS

In order to prove the effectiveness and robustness of the proposed method, we present some results obtained on simulated phase and amplitude data. Results with real data will be presented in the final version of the paper. We consider in figure 2-(a) a height profile (64×64 pixels) with a maximum height of $120m$ exhibiting both smooth and discontinuous areas. A corresponding amplitude image of the same scene is shown in the figure 2-(b). We used a total of 8 frequencies in the $[5GHz, 9GHz]$ interval to generate independent interferograms and we added interferometric noise with two different region coherences ($\gamma_1 = 0.8, \gamma_2 = 0.4$). The low coherence region is at the bottom right corner of the presented profile. In figure 2-(c), we show the $5GHz$ noisy interferogram. It is important to note that the profile is ambiguous for all the working frequencies. In fact, there are phase jumps of more than π that violate the Itoh condition. For this reason, a classical single frequency phase unwrapping method would fail and the multichannel approach can overcome this problem. In figure 2-(d), we show the noisy amplitude corrupted by the speckle. Results of height reconstruction with the multichannel phase unwrapping approach proposed in [1], with 4 channels (resp. 8 channels), are presented in figure 2-(e) (resp. figure 2-(f)) and the root mean squared reconstruction errors are respectively 0.37 and 0.12. The 3D reconstruction with the proposed approach is presented in figure 2-(g), where the root mean squared reconstruction error is 0.10. We see clearly the contribution of the proposed approach to restore profiles

in presence of low coherent interferogram regions and high discontinuities. The approach in [1] requires more channels and an accurate estimation of the regularization hyperparameter to correctly unwrap the phase. Otherwise, regions with very low coherence and high discontinuities could not be reconstructed correctly as we see in figure 2-(e). The proposed approach, instead, requires only a few number of channels to correctly unwrap and regularize the phase, since an additional discontinuity information is provided by the amplitude data. The restored amplitude is shown in figure 2-(h). We can notice the effectiveness of the proposed prior when dealing with discontinuities:the obtained reconstructions preserve the discontinuities in both original phase and amplitude images.
Concerning complexity of our algorithm, since the optimization process is based on graph-cut technique, where max-flow is computed in polynomial time, our approach is optimal both in time computation and memory occupation. For the presented experiment, the time needed by our algorithm to converge toward a good local minimum is $2(Mn) : 40(Sec)$, and the memory used for graph construction is equivalent to the size of the restored profile image.

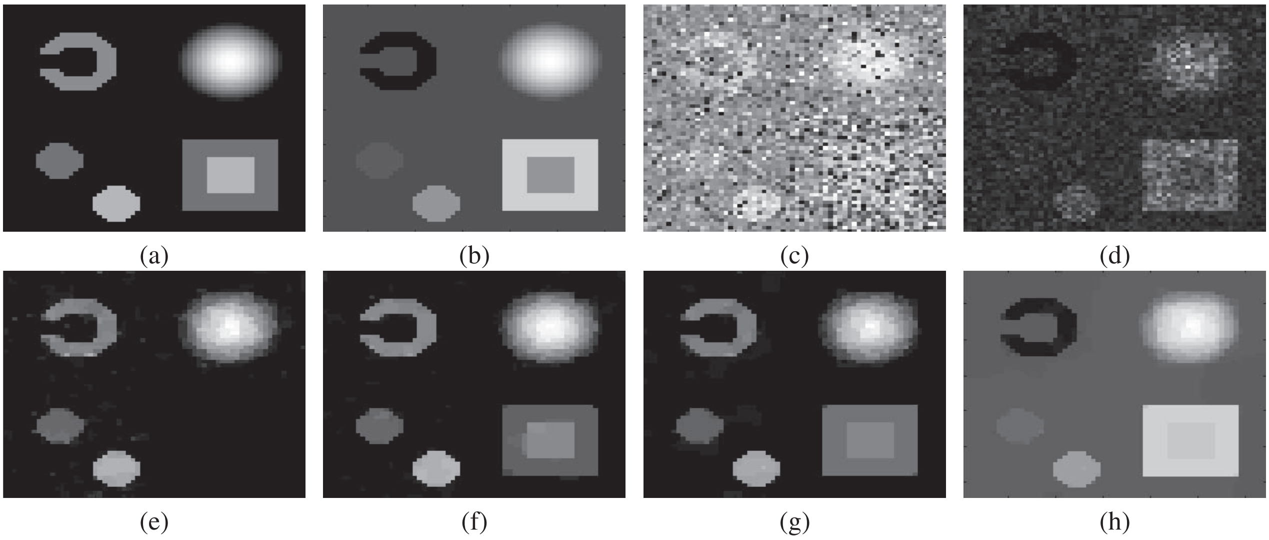


Fig. 2. Reconstruction results: (a)True profile, (b)True amplitude, (c)Interferogram at $f = 5Ghz$ with two region coherences (γ_1, γ_2), (d)Noisy amplitude, (e) Profile reconstruction with 4 frequencies using [1], (f) Profile reconstruction with 8 frequencies using [1], (g) and (h) Profile reconstruction and amplitude regularization with 4 frequencies using the proposed approach.

5. CONCLUSIONS

In this work, we developed a new Markovian 3D urban area reconstruction approach based on a joint multichannel phase and amplitude model. A graph-cut based optimization algorithm is proposed to solve the energy minimization problem. We tested this approach on simulated data and we obtained good results both in term of reconstruction error and computational time. Obtained results also proved the effictiveness of the approach when dealing with strong discontinuities and low coherences areas.

6. REFERENCES

[1] G. Ferraioli, A. Shabou, F. Tupin, and V. Pascazio, “Multichannel phase unwrapping with graph cuts,” *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 3, pp. 562–566, 2009.

[2] L. Denis, F. Tupin, J. Darbon, and M. Sigelle, “Joint regularization of phase and amplitude of InSAR data: Application to 3D reconstruction,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3774–3785, 2009.

[3] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.