HAL
open science

Urban surface recovery through graph-cuts over SAR tomographic reconstruction

Clément Rambour, Loïc Denis, Florence Tupin, Hélène Oriot

To cite this version:

Clément Rambour, Loïc Denis, Florence Tupin, Hélène Oriot. Urban surface recovery through graph-cuts over SAR tomographic reconstruction. 2019 Joint Urban Remote Sensing Event (JURSE), May 2019, Vannes, France. 10.1109/JURSE.2019.8809001 . hal-02924933

HAL Id: hal-02924933

https://hal.science/hal-02924933v1

Submitted on 28 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Urban surface recovery through graph-cuts over SAR tomographic reconstruction

1st Clément Rambour *LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France* 2nd Loïc Denis *Univ Lyon, UJM-Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien, UMR 5516, F-42023, SAINT-ETIENNE, France* 3rd Florence Tupin *LTCI, Télécom ParisTech, Université Paris-Saclay, 75013 Paris, France* 4th Hélène Oriot *ONERA, The French Aerospace Laboratory, 91761 Palaiseau, France*

***Abstract*—The understanding of tomograms can be a difficult task when the observed scene is complex. The Synthetic Aperture RADAR (SAR) ranging acquisition induces geometrical distortions depending on the local slope of the back-scattering surface. Over dense urban areas, this particular geometry produces tilted facades instead of straight buildings. Moreover the high dynamic in SAR images of urban areas may lead to 3-D representation with very bright voxels eclipsing other echoes. In this paper we present a graph-cut approach adapted to any tomographic reconstruction technique to segment the urban surface in a 3-D tomographic reconstruction. Results on real data from a stack of 40 TerraSAR-X images are presented.**

***Index Terms*—SAR, tomography, graph-cuts, urban areas**

Identify applicable funding agency here. If none, delete this.

I. Introduction

SAR tomography has been successfully used for different kind of applications over the last two decades. Experiences over uniform environments such as forest or glaciers [1]–[3] have produced 3-D models of the scatterers density of reflectivity. With the increase of resolution SAR tomography has been also experienced over urban areas allowing to have global rendering of observed cities.

The SAR tomography is basically the construction of a synthetic antenna in the elevation direction, *i.e.* orthogonal to the line of sight. This new antenna is formed by exploiting the different acquisitions available over a given scene. The focusing in the elevation direction can theoretically be done using a discrete inverse Fourier

transform of the data in the stacking direction.

To improve the resolution, different super-resolution approaches have been proposed. The Compress Sensing (CS) technique [4], [5] based on ℓ_1-norm minimization of the back-projection of the stack has shown very promising results for dense urban areas. One recent CS-like approach [6] proposes to perform the back-projection of a 3-D cube corresponding to many pixels at once. This last method also performs the inversion directly in ground geometry to be able to use geometrical priors (*cf.* Fig. 1).

Having the reconstruction expressed in ground geometry is first step to help a potential viewer locating typical urban structures such as walls, ground and rooftops. The concentration of scatterers, secondary lobes and outliers however still produce ambiguities in the interpretation. In this paper, we present a new graph-cut approach performing the segmentation of the urban scene from a tomographic reconstruction. The proposed method favors smooth reconstructions along the (x, y, z) axis that are faithful to the tomograms. The developed method is very general and could be used with any 3-D reconstruction in input as long as it corresponds to a physical representation of the signal (*e.g.* its amplitude or power). Here only the 3-D inversion under a sparse prior is used to illustrate the method due to its good performances on urban areas and as it performs the inversion in the desired geometry. Finally we present an algorithm that alternates between a 3-D reconstruction and a surface segmentation step. Through the iterations, the segmentation information is used to improve the sparsity priors.

II. Ground frame 3-D back-projection

The SAR tomographic stack defines a set of N Single Look Complex (SLC) SAR images. After co-registration and phase calibration processing, the complex value for a pixel in the nth image located at the position (x, r) is

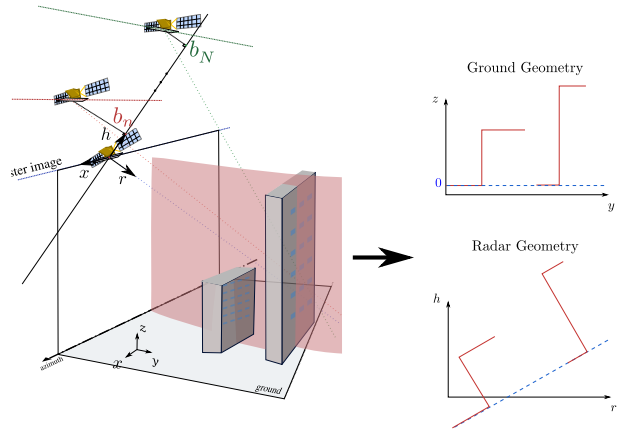

Fig. 1. Geometry of acquisition

given by the integration of the complex reflectivity u in the corresponding radar cell:

$$\underline{v}_n(x, r) = \iint_{(y,z)\in\Delta_r} u(x, y, z)\exp\big(-j\xi_n z\big)\mathrm{d}y\mathrm{d}z + \epsilon \tag{1}$$

The variables x, y and z denote the coordinates in the ground geometry. For sake of clarity the azimuth and x axis are chosen the same. The random variable $\epsilon(x, r)$ models the additive Gaussian thermal noise. The r-th radar resolution cell is defined by: $\Delta_r = \{\ (y, z)\ |\ r - \delta_{\text{range}}/2 \leq \rho_{y,z} \leq r + \delta_{\text{range}}/2\ \}$, with δ_{range} the step in range direction. The parameter $\xi_n = \frac{4\pi b_n}{\lambda r \sin\theta}$ is the spatial impulse associated to the sampling of the scene for each baseline, b_n is the n-th baseline, θ the incidence angle of the master sensor and λ the radar wavelength.

For urban areas, CS has been successfully used to inverse the previous equation under a sparsity prior. It has proved to achieve super resolution in scatterer unmixing in dense urban areas [4], [5]. Similar to conventional CS but able to process the estimation directly in ground geometry the 3-D inversion proposed in [6] allows to use geometrical priors. Here we will only use a sparse prior and thus the estimation of the reflectivity for all the scene is the solution of the problem:

$$\hat{\boldsymbol{u}} = \underset{\boldsymbol{u}}{\text{argmin}} \ ||\boldsymbol{\Phi}\boldsymbol{u} - \boldsymbol{v}||_2^2 + \mu||\boldsymbol{u}||_1 \tag{2}$$

The matrix $\boldsymbol{\Phi}$ is the back-projection operator that projects a scene expressed in ground geometry in the SAR tomographic stack:

$$\underline{\boldsymbol{\Phi}}_{k,l+(n-1).N} = \begin{cases} 0 \text{ if } \rho_{1;y_l\, z_l} \notin [r_k - \frac{\delta_{\text{range}}}{2}, r_k + \frac{\delta_{\text{range}}}{2}] \\ \exp(-j\xi_n z_l) \text{ otherwise,} \end{cases} \tag{3}$$

The parameter μ balances the importance of the sparsity prior with respect to the data fidelity. The tuning of this parameter is a key point to have a good estimation of the reflectivity. However the knowledge of the additive noise level is hardly enough to get rid of outliers and/or side-lobes. As no prior information on the number of scatterers in each cell is available this parameter is hard to set. In the next part we present the graph-cuts based segmentation of the urban surface which we will use to refine the tomographic reconstruction.

III. Graph construction and urban surface segmentation

From a 3-D estimation of the reflectivity, we seek for the back-scattering urban surface accordingly to the following hypotheses:

1) The surface is located near the brightest voxel in the reconstruction.
2) A ray from the sensor to the intersection of a radar cell with the ground goes through the surface exactly once.
3) The urban surface is smooth and piecewise uniform either along the z direction or alongside the (x, y) plan.
4) The walls are straight and vertical.

To answer this segmentation problem we need to design a cost function that increases as these hypotheses are violated. From the first and second assumptions, for each ray the cost function has to increase with the size of the gap between the surface intersection and the brightest voxel on a ray. When this criterion is minimized the obtained surface is totally faithful to the original reconstruction. Hypothesis 3 and 4 help do deal with potential holes in the reconstruction or lobe spreading along the elevation direction.

As the tomographic reconstructions are subjects to outliers, rather than estimating for each ray the maximum of intensity that may be ambiguous, we use the cumulative reflectivity functions defined as:

$$C^-(r_s) = \int_{r_{\min}}^{r_s} |u(r)|\mathrm{d}r \quad \text{and} \quad C^+(r_s) = \int_{r_s}^{r_{\max}} |u(r)|\mathrm{d}r \tag{4}$$

with r_s being the range of the surface along the ray. If the surface cuts the ray such as $C^-(r_s) < C^+(r_s)$ then too much reflectivity is left "under" it from the sensor point of view. On the contrary having $C^-(r_s) > C^+(r_s)$ means that most of the bright voxels are located before the surface. The quantity $C^-(r_s) - C^+(r_s)$ is then an indication on whether the surface is close to the median of the reflectivity along the ray. The penalty we use for the surface location is then given by:

$$D(r) = \int_{r_{\min}}^{r} \left[C^-(r_s) - C^+(r_s)\right]_+ \mathrm{d}r_s + \int_{r}^{r_{\max}} \left[C^+(r_s) - C^-(r_s)\right]_+ \mathrm{d}r_s\,, \tag{5}$$

with $[.]_+$ being the Heaviside function. The only value for which the two terms in equation (5) are both zero correspond to $r = r_{\text{equi}}$ the median of the reflectivity along the ray. Moreover only one of the term is non-zero as when $r < r_{\text{equi}}$, $C^-(r_s) < C^+(r_s)$ and then $[C^-(r_s) - C^+(r_s)]_+ = 0$ and inversely when $r > r_{\text{equi}}$. The penalty function D is then a monotonous increasing function of the distance $|r - r_{\text{equi}}|$. Finally the data

penalty is the sum of the penalization D for all the possible rays going through the surface $\mathscr{S}$:

$$P_d(\mathscr{S}) = \int_{\tau \in \mathscr{R}} D_\tau(r_{\tau \to \mathscr{S}})\, \mathrm{d}\tau \tag{6}$$

where $r_{\tau \to \mathscr{S}}$ is the range position of the surface for the ray τ and D_τ is the penalty function associated to this ray.

To satisfy the third hypothesis and prevent the surface to present to much oscillations, we apply a ℓ_1 penalization on its variations. We define the spatial penalty as:

$$P_s(\mathscr{S}) = \iiint |\nabla \mathscr{S}(x, y, z)|\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \tag{7}$$

with $\nabla \mathscr{S}(x, y, z)$ being the gradient of the surface. The estimated surface is then the solution of the following minimization problem:

$$\hat{\mathscr{S}} = \min_{\mathcal{S}} P_d(\mathscr{S}) + \beta P_s(\mathscr{S}) \tag{8}$$

The parameter β balances the fidelity to the 3-D tomographic reconstruction and the spatial smoothness of the surface.

After discretization of the surface and the set of rays, the problem (8) can be equivalently formulated as a maximum flow / minimum cut problem for a well designed graph as presented in Fig. 2. Each node in the graph corresponds to a voxel in the reconstruction plus the source s and the sink t. The capacities of the edges linking the node i to s and t are set respectively to $[C^-(r_i) - C^+(r_i)]_+$ and $[C^+(r_i) - C^-(r_i)]_+$, with r_i being the distance from the sensor to the node i. The capacities of two neighboring nodes is set equal to β. Note that it is necessary to have the volume expressed in ground geometry in order to favor the surface smoothness along either z or (x, y). The minimum cut of this 3-D graph corresponds then to the surface separating the less costly set of edges and corresponds to the minimization of (8) up to the discretization of the scene.

We still have to address the last assumption of straight walls. To avoid having snaking walls in the final estimation, we add infinite capacity edges going from one node to its upper neighbor along the z direction. As only the capacities of the cut edges going from the source to the sink are counted in the cut, this penalty prevents the surface to present crenelation shape along a vertical line. The construction of the graph and the computation of the minimum cut is done using Boykov and Kolmogorov graph-cuts library [7].

IV. Alternated 3-D Reconstruction and Surface Segmentation (REDRESS)

As stated in the introduction the sparsity parameter is difficult to tune locally with CS methods. Here we propose to use the 3-D knowledge provided by the estimation of the urban surface to set the parameter μ locally in the 3-D space. The tomographic reconstruction and surface segmentation steps can then be repeated to have refined estimation of the scene. For the kth iteration the parameter $\boldsymbol{\mu}_k \in \mathbb{R}^{N_x . N_y . N_z}$ is computed as follows:

$$\boldsymbol{\mu}_k(x, y, z; \mathscr{S}) = \mu_0 + \frac{b}{(n-1)^2} \left(\frac{k}{n-k}\, d((x, y, z), \mathscr{S}) \right)^2 \tag{9}$$

where $d((x, y, z), \mathscr{S})$ is the Euclidean distance from the point $\boldsymbol{p} = (x, y, z)^T$ to the surface and n the total number of iterations. μ_0 is the first global value for the parameter and is the one applied to points that lie on the surface. To avoid penalizing too much points close to the surface the square of the distance is used with an increasing coefficient between 0 to 1. b corresponds to the final coefficient applied to the square of the distance and balances how fast the penalty increases through the iterations. The REDRESS algorithm is then described by:

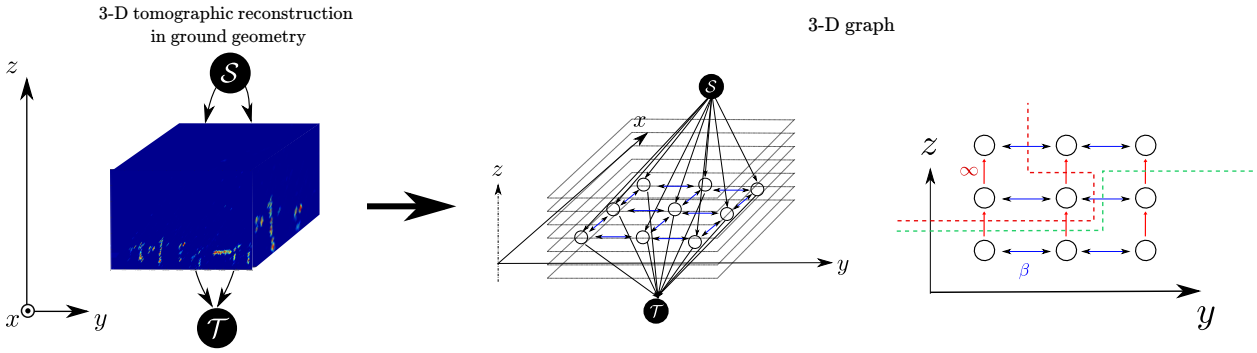

Fig. 2. Construction of the graph from the 3-D tomographic estimation of the reflectivity in ground coordinate. Each voxel is associated to a node and is linked to its direct neighbors plus the source s and the sink t. The blue edges have the capacity β and the red ones a infinite capacity. On the right, two examples of a cut: a feasible one in green and an one with a infinite cost in red as an infinite capacity edge is going from s to t.

while $k < n$ **do**

$\hat{\underline{\boldsymbol{u}}} \leftarrow$ `3-D Inversion`$(\underline{\boldsymbol{v}}, \boldsymbol{\mu}(\mathscr{S}))$

$\mathscr{S} \leftarrow$ `graph cut`$(\hat{\underline{\boldsymbol{u}}})$

$k \leftarrow k + 1$

end while

V. Results

The performances of the proposed algorithm are demonstrated on a stack of 40 TerraSAR-X images of the city of Paris. The baselines are non-uniformly sampled in space and time. The baselines, the mean intensity SAR image and the optical view of the area are shown Fig. 3.

The result obtained using the REDRESS algorithm as well as the ground truth extracted from Google Earth © are presented in Fig. 4. Five iterations were used for the algorithm. The observed building is well represented with correct height. Even details such as the oscillating structure of the rooftop of pillars are visible in the estimation. In the last version of this paper, results on a bigger area will be presented.

VI. Conclusion

In this paper we present a new graph-cut based urban surface segmentation method. This segmentation

Fig. 3. Data used for the inversion and surface segmentation. The building seen correspond to the ministry of foreign affair in Paris. From top to bottom, temporal intensity mean of the SAR pile, optical image of the areas and baselines used.

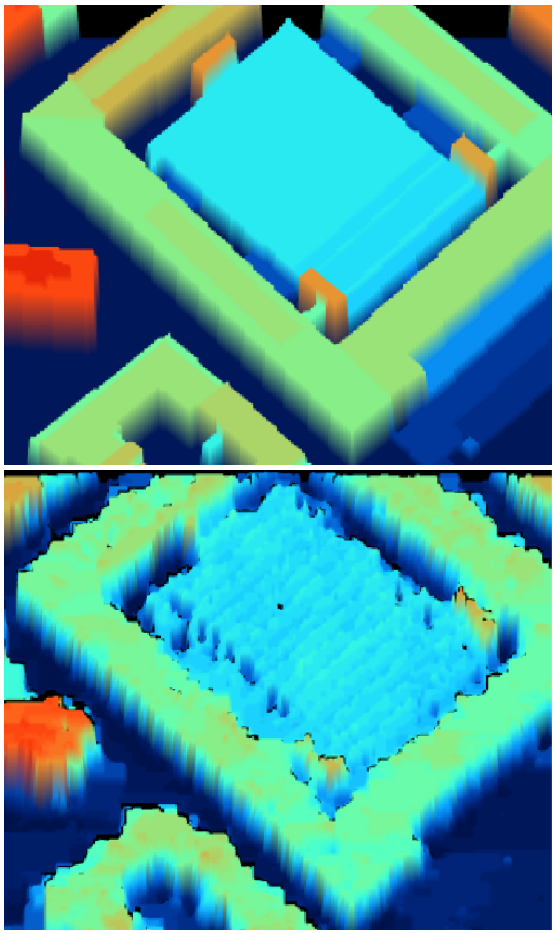

Fig. 4. On top ground truth, on bottom surface estimation for the observed building

technique add a significant amount of information in the interpretation of the 3-D data. This knowledge can then be used to answer to the recurrent problem of parameter tuning in sparse tomographic SAR reconstruction. The algorithm REDRESS alternates several steps of tomographic estimation and segmentation to refine both the sparse parameter tuning and the estimated surface. Results on real data for dense urban landscape led to interesting detailed results.

REFERENCES

[1] Y. Huang, L. Ferro-Famil, and A. Reigber, "Under foliage object imaging using SAR tomography and polarimetric spectral estimators," in *8th European Conference on Synthetic Aperture Radar*, June 2010, pp. 1–4.

[2] S. Tebaldini and F. Rocca, "Multibaseline Polarimetric SAR Tomography of a Boreal Forest at P- and L-Bands," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 232–246, Jan 2012.

[3] S. Tebaldini, T. Nagler, H. Rott, and A. Heilig, "Imaging the Internal Structure of an Alpine Glacier via L-Band Airborne SAR Tomography," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7197–7209, Dec 2016.

[4] X. X. Zhu and R. Bamler, "Tomographic SAR Inversion by L_1 -Norm Regularization ; The Compressive Sensing Approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, Oct 2010.

[5] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-Dimensional SAR Focusing From Multipass Signals Using Compressive Sampling," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 488–499, Jan 2011.

[6] C. Rambour, L. Denis, F. Tupin, J.-M. Nicolas, and H. Oriot, "SAR Tomography Of Urban Areas: 3D Regularized Inversion In The Scene Geometry," in *2018 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2018.

[7] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE transactions on pattern analysis and machine intelligence*, vol. 26, no. 9, pp. 1124–1137, 2004.