

SAR TOMOGRAPHY OF URBAN AREAS: 3D REGULARIZED INVERSION IN THE SCENE GEOMETRY

Clément Rambour, Loïc Denis, Florence Tupin, Jean-Marie Nicolas, Hélène Oriot

▶ **To cite this version:**

Clément Rambour, Loïc Denis, Florence Tupin, Jean-Marie Nicolas, Hélène Oriot. SAR TOMOGRAPHY OF URBAN AREAS: 3D REGULARIZED INVERSION IN THE SCENE GEOMETRY. 2018 IEEE International Geoscience and Remote Sensing Symposium, Jul 2018, Valence, Spain. hal-01873659

HAL Id: hal-01873659
https://hal.science/hal-01873659

Submitted on 13 Sep 2018

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.

SAR TOMOGRAPHY OF URBAN AREAS: 3D REGULARIZED INVERSION IN THE SCENE GEOMETRY

Clément Rambour[1]*, Loïc Denis*[2]*, Florence Tupin*[1]*,Jean-Marie Nicolas*[1]*, Hélène Oriot*[3]*,*

[1]LTCI, Télécom ParisTech, Université Paris Saclay, Paris, France
[2] Univ Lyon, UJM- Télécom Saint-Etienne, CNRS, Institut d'Optique Graduate School, Laboratoire Hubert Curien UMR 5516, F-42023, SAINT-ETIENNE, France
[3]ONERA, The French Aerospace Lab, F-91761 - Palaiseau, France

ABSTRACT

Starting from a stack of co-registered SAR images in interferometric configuration, SAR tomography performs a reconstruction of the reflectivity of scatterers in 3-D. Scatterers seen within the same resolution cell in each SAR image can be separated by jointly unmixing the SAR complex amplitude observed throughout the stack. In urban areas, Compress Sensing (CS) approaches have been applied to achieve super-resolution in the estimation of the position of the scatterers. However, even if all the local information coming from a stack at a given pixel is used, the structural information that is inherent to the image is not directly used to improve the rendering of the scene. This paper addresses the problem of adding structural constraints to sparse tomographic reconstructions of urban areas. We derive an algorithm allowing the inversion of tomographic data under structural constraints and illustrate its performances on a stack of Spotlight TerraSAR-X images.

Index Terms— SAR tomography, structural information, spatial regularization

1. INTRODUCTION

SAR tomography is the extension of the 2-D SAR imaging to three dimensions. As conventional 2-D SAR imaging uses a synthetic aperture in azimuth direction, 3-D SAR imaging is performed by a synthetic aperture in the elevation direction by collecting several images from parallel tracks. Using multi-baseline interferometry techniques on the well-calibrated SAR images stack, it is possible to retrieve the localization of the scatterers in the third dimension. This approach allows separating scatterers mapped in the same resolution cell, which is likely to happen on the dense urban area due to the layover phenomenon.

To perform 3-D SAR focusing, application of the inverse discrete Fourier transform gives a resolution inversely proportional to the synthetic baseline aperture in the cross-range direction. Depending on the baseline repartition this resolution is generally much worse compared to the resolution in azimuth or range. Several spectral super-resolution techniques have been used to overcome this limitation [1] but they all rely on the covariance matrix. This matrix has to be estimated which adds a strong limitation on the number of images that can be used and can be difficult when the observed scene is very heterogeneous. Compress Sensing (CS) also leads to super-resolution but by performing a direct inversion of the data under sparsity constraint. As it doesn't require the estimation of the covariance matrix, this method is far more robust for unmixing multiple close point scatterers.

CS provides a sparse reconstruction of the 3-D scene [2] but is applied pixel by pixel without exploiting any structural knowledge of the scene. Different strategies have been developed to re-process the estimated signal in order to improve its spatial coherence [3], but until now the structural information has not been taken into account directly in the inversion step.

Here, we develop an algorithm that allows to inverse a tomographic stack to retrieve a sparse 3-D complex scene using structural geometric priors. To do so, we first write the transformation that allows us to express the scene in ground geometry and then we describe the inversion process. We present 3-D reconstructions obtained from a dataset of TerraSAR-X images over the city of Paris.

2. GROUND TO RADAR GEOMETRY

A SAR tomographic stack consists of a collection of N co-registered SAR images. The geometry of the objects and the sensors is illustrated Fig. 1. In the absence of moving objects, the value of the complex pixel in image n at position (x, r) is [4]

$$\underline{v}_n(x,r) = \int\limits_{(y,z)\in\Delta(x,r)} \underline{u}(x,y,z)\exp\big(-\underline{j}\xi_n h(y,z)\big)\mathrm{d}y\mathrm{d}z + \underline{n} \tag{1}$$

This project has been funded by ANR (the French National Research Agency) and DGA (Direction Générale de l'Armement) under ALYS project ANR-15-ASTR-0002

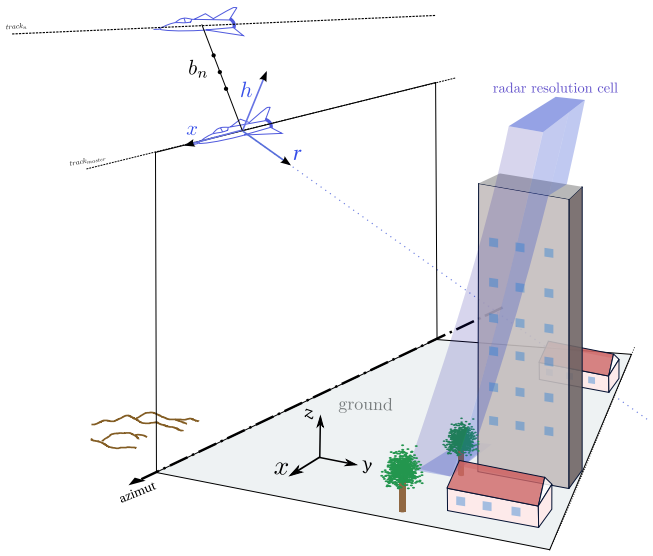


Fig. 1. **Radar and ground geometry:** The geometry of the sensors (x, r, h) doesn't allow a good expression of the structural particularity of the scene. In particular, it is not natural to express the orthogonality of ground / walls or walls / roofs. We place ourselves in the ground geometry (x, y, z) to be able to take those observations into account.

Complex variables are from now on underlined. Here, $\Delta(x, r)$ is the extension of the radar resolution cell along the elevation direction h, $\underline{u}(x, y, z)$ the complex reflectivity of the scatterer located at position (x, y, z) and $h(y, z)$ its elevation. We chose to have the azimut and x axis being the same for simplification. $\xi_n = \frac{4\pi b_n}{\lambda r}$ is the spatial frequency depending on the baseline b_n, λ stand for the wavelength and θ is the incidence angle of the master antenna. n is a complex additive white noise coming from the sensors.

As we can see the ground geometry that describes best the objects is not the same as the radar geometry which will depend on the incidence angle. Usually, the tomographic inversion is processed using the radar geometry, but in order to express some structural constraints we will use the ground geometry during the inversion process. We now refer to $\underline{\boldsymbol{v}} \in \mathbb{C}^{N_x.N_r.N}$ as the tomographic stack and $\underline{\boldsymbol{u}} \in \mathbb{C}^{N_x.N_y.N_z}$ as the 3-D unknown scene, with N_r, N_x, N_y and N_z being the number of bins in the corresponding directions.

We write $\underline{\boldsymbol{\Phi}} \in \mathbb{C}^{(N_x.N_r.N)\times(N_x.N_y.N_z)}$ the matrix that express the signal back-scattered from a scene sampled in ground geometry to the antennas. An element of $\underline{\boldsymbol{\Phi}}$ is then

$$\underline{\boldsymbol{\Phi}}_{i,j+(n-1).N} = \delta_{r_i,c(y_j,z_j)} \exp\left[-\underline{j}\xi_n h(y_j, z_j)\right] \quad (2)$$

The function $c(y, z)$ return the corresponding range of a point located at the position (y, z). The matrix $\underline{\boldsymbol{\Phi}}$ is highly sparse which can be used to speed-up the computation and its non-null elements correspond to permutations of the classical sensing matrix used in SAR tomography. We now have designed a linear operator that allows us to express the tomographic stack from the 3-D complex scene :

$$\underline{\boldsymbol{v}} = \underline{\boldsymbol{\Phi}}\underline{\boldsymbol{u}} + \underline{\boldsymbol{n}} \quad (3)$$

3. TOMOGRAPHIC INVERSION

In this part we will derive the inversion of equation (3) using some basic priors on the urban topography. To achieve super resolution, we will have a number of unknowns variables far more important than the size of $\underline{\boldsymbol{v}}$. To be able to estimate $\underline{\boldsymbol{u}}$ we will add two constraints. First, we will assume that the 3-D scene is sparse which is the same hypothesis as in CS. Secondly, we will make the assumption that the objects are smooth in the x, y and z directions. Those hypothesis express the fact walls and roofs are seen as punctual objects in at least one direction and are smooth in the other two directions. We then want to solve the following minimization problem

$$\begin{aligned} \hat{\underline{\boldsymbol{u}}} &= \underset{\underline{\boldsymbol{u}}}{\operatorname{argmin}} \, ||\underline{\boldsymbol{\Phi}}\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}||_2^2 + \mathcal{R}(|\underline{\boldsymbol{u}}|) \\ \text{with} \quad \mathcal{R}(|\underline{\boldsymbol{u}}|) &= \mu_x \cdot \|\boldsymbol{D}_x|\underline{\boldsymbol{u}}|\|_2^2 + \mu_y \cdot \|\boldsymbol{D}_y|\underline{\boldsymbol{u}}|\|_2^2 \\ &\quad + \mu_z \cdot \|\boldsymbol{D}_z|\underline{\boldsymbol{u}}|\|_2^2 + \mu_{L_1} \cdot \||\underline{\boldsymbol{u}}|\|_1 \end{aligned} \quad (4)$$

μ_x, μ_y, μ_z and μ_{ℓ_1} are the parameters allowing to tune the impact of each term in $\mathcal{R}$. Here the regularization is applied to the modulus of the 3-D reflectivity. One big reason is that we don't want to risk to converge to zero as we smooth two neighboring scatterers with opposite phase. The introduction of the modulus add a lot of complexity to this problem even with a relatively simple prior. To be able to solve (4), we introduce two more variables and chose an alternating minimization scheme. We now want to solve

$$\begin{aligned} \hat{\underline{\boldsymbol{u}}} &= \underset{\underline{\boldsymbol{u}}}{\operatorname{argmin}} \quad \|\underline{\boldsymbol{\Phi}}\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}\|_2^2 + \mathcal{R}(\boldsymbol{w}) \\ \text{s.t.} \quad \underline{\boldsymbol{u}} &= \underline{\boldsymbol{f}} \\ |\underline{\boldsymbol{f}}| &= \boldsymbol{w} \end{aligned} \quad (5)$$

The real image $\boldsymbol{w}$ will converge to the modulus of the image $\underline{\boldsymbol{u}}$ and the complex image $\underline{\boldsymbol{f}}$ links them. We solve (5) using a modified ADMM [5] algorithm by jointly optimizing over $\underline{\boldsymbol{u}}$ and $\boldsymbol{w}$ and updating the two Lagrange multipliers :

$$\begin{aligned} &\bullet \quad \{\hat{\underline{\boldsymbol{u}}}, \hat{\boldsymbol{w}}\} \leftarrow \underset{\underline{\boldsymbol{u}}, \boldsymbol{w}}{\arg\min} \Big\{ \frac{1}{2}||\underline{\boldsymbol{\Phi}}\underline{\boldsymbol{u}} - \underline{\boldsymbol{v}}||_2^2 \\ &\qquad + \frac{\beta_1}{2}||\underline{\boldsymbol{f}}(\underline{\boldsymbol{u}}, \boldsymbol{w}) - \underline{\boldsymbol{u}} + \hat{\underline{\boldsymbol{d}}}_1||_2^2 \\ &\qquad + \frac{\beta_2}{2}\big\|\boldsymbol{w} - |\underline{\boldsymbol{f}}(\underline{\boldsymbol{u}}, \boldsymbol{w})| + \hat{\boldsymbol{d}}_2\big\|_2^2 + \mathcal{R}(\boldsymbol{w}) \Big\} \\ &\bullet \quad \hat{\boldsymbol{d}}_2 \leftarrow \hat{\boldsymbol{d}}_2 + \hat{\boldsymbol{w}} - |\underline{\boldsymbol{f}}(\hat{\underline{\boldsymbol{u}}}, \hat{\boldsymbol{w}})| \\ &\bullet \quad \hat{\underline{\boldsymbol{d}}}_1 \leftarrow \hat{\underline{\boldsymbol{d}}}_1 + \underline{\boldsymbol{f}}(\hat{\underline{\boldsymbol{u}}}, \hat{\boldsymbol{w}}) - \hat{\underline{\boldsymbol{u}}}\,. \end{aligned} \quad (6)$$

β_1 and β_2 are two parameters needed for the optimization and $\underline{\boldsymbol{f}}(\underline{\boldsymbol{u}}, \boldsymbol{w})$ is given by

$$\underline{\boldsymbol{f}}(\underline{\boldsymbol{u}}, \boldsymbol{w}) = \left[\frac{\beta_1 \cdot |\underline{\boldsymbol{u}} - \underline{\hat{\boldsymbol{d}}}_1| + \beta_2 \cdot (\boldsymbol{w} + \hat{\boldsymbol{d}}_2)}{\beta_1 + \beta_2}\right]^+ \cdot \exp\left[\mathrm{i} \cdot \arg\left(\underline{\boldsymbol{u}} - \underline{\hat{\boldsymbol{d}}}_1\right)\right] \quad (7)$$

The first step in (6) is easy to optimize as it is a sum of quadratic and linear functions. As the modulus is positive, the ℓ_1 norm can be seen as the sum of all the elements in the vector. The algorithm described in (6) gives a new tool capable to inverse a tomographic problem under structural priors. Even if the model we choose is still simple, it gives good results on urban areas as shown in the next section. As the framework is completely independent of the regularization chosen, the structure of the algorithm can be used to add even more sophisticated priors on the reconstructed data.

4. EXPERIMENTAL RESULT

The algorithm described in the previous section has been implemented and tested on a dataset of 40 TerraSAR-X images over the *front de Seine* in the south-west of Paris, France. The observed scene is presented Fig. 3[1].The spatial (orthogonal component) and temporal baseline are shown in Fig. 2. The total spatial baseline span is more than 775m and the total temporal baseline more than 5 years with a large gap of almost two years in it.
Fig. 4 we present the intensity of the 3-D inversion over the Mirabeau Tower and we also plot in Google Hearth the height of the discretized reconstruction obtained by thresholding the intensity. Fig. 5 we present the intensity of the 3-D inversion over the Ministry of Foreign Affairs and the height of the discretized reconstruction in Google Hearth.

5. CONCLUSION

In this paper, we have proposed a new algorithm for the inversion of a stack of tomographic images. In this approach, we add structural constraints on the modulus of the 3-D reflectivity in order to describe more accurately a dense urban scene. This algorithm has been implemented and gives good results on real images. Finally, as the priors we add can be changed, we introduce a framework capable of performing 3-D tomographic inversion under spatial constraint.

[1]These data have been provided by the DLR in the framework of the project LAN1746

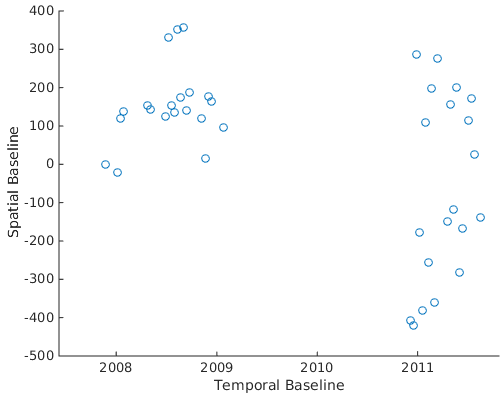


Fig. 2. Spatial Baseline (orthogonal part) vs. Temporal Baseline

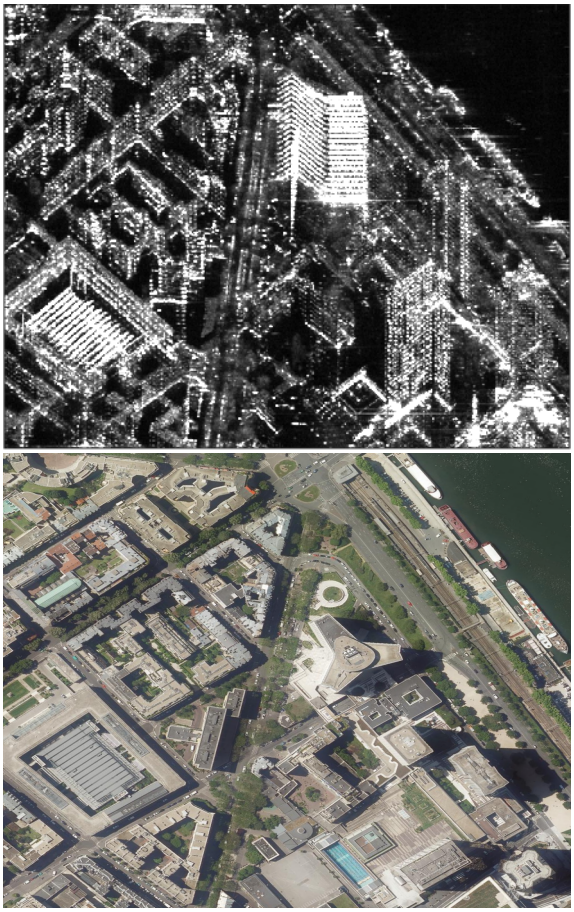


Fig. 3. (Top) TerraSAR-X intensity image of the observed scene (Paris). The building back-scattering a strong signal in the center af the image is the Mirabeau Tower. In the bottom left corner is the Ministry of Foreign Affairs with a periodic appearing rooftop. (Bottom) Optical image from Geoportail of the scene.

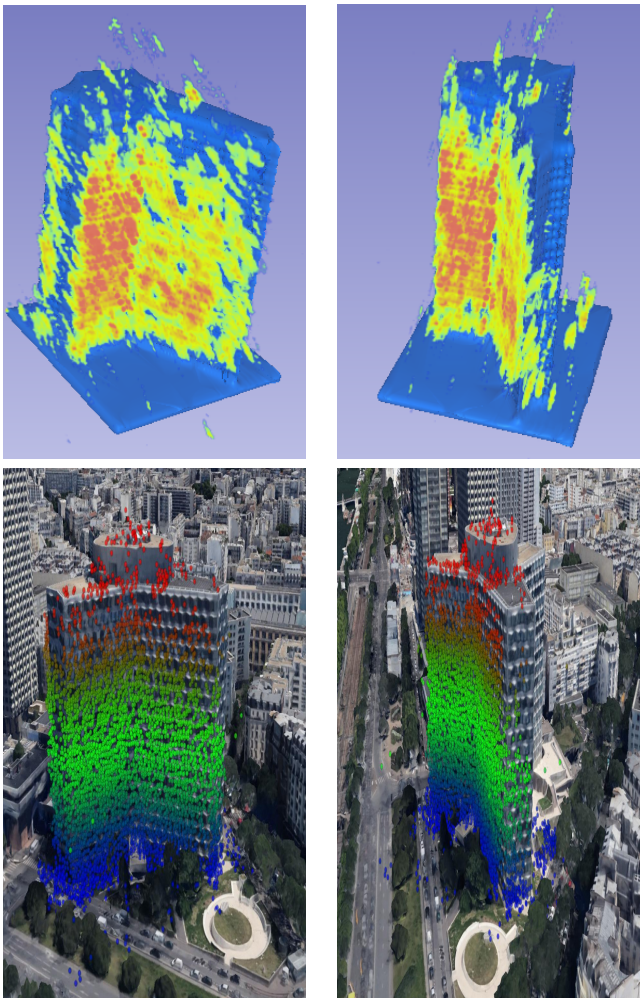


Fig. 4. (top) Intensity of the 3-D reconstruction of the Mirabeau Tower [6]. (bottom) height of the brightest points in the 3-D reconstruction plotted in Google Hearth.

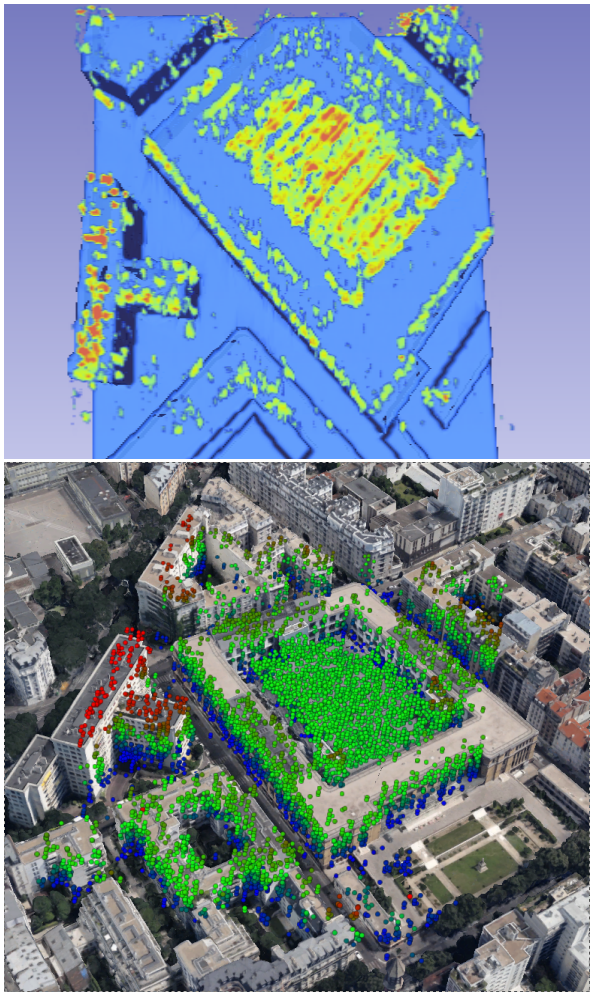


Fig. 5. (top) Intensity of the 3-D reconstruction of the Ministry of Foreign Affairs [6]. (bottom) height of the brightest points in the 3-D reconstruction plotted in Google Hearth.

6. REFERENCES

[1] R. Moses P. Stoica, ," in *Introduction to Spectral Analysis*. Prentice Hall, 1997.

[2] X. X. Zhu and R. Bamler, "Tomographic sar inversion by l_1 -norm regularization 2014;the compressive sensing approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, Oct 2010.

[3] X. X. Zhu and M. Shahzad, "Facade reconstruction using multiview spaceborne tomosar point clouds," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3541–3552, June 2014.

[4] G. Fornaro, F. Lombardini, and F. Serafino, "Three-dimensional multipass sar focusing: experiments with long-term spaceborne data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 702–714, April 2005.

[5] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[6] Andriy Fedorov et al., "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magnetic Resonance Imaging*, vol. 30, no. 9, pp. 1323 – 1341, 2012, Quantitative Imaging in Cancer.