DOI 10.1109/IGARSS.2017.8128340 – http://ieeexplore.ieee.org/document/8128340/

REGULARIZATION AND COMPLETION OF TOMOSAR POINT CLOUDS IN A PROJECTED HEIGHT MAP DOMAIN

Andreas Ley, Olivier D'Hondt and Olaf Hellwich

Technische Universität Berlin, Germany. e-mail: name.surname@tu-berlin.de

ABSTRACT

TomoSAR (Tomographic SAR) is a technique allowing to extend SAR imaging to the third dimension by using several images of a scene acquired from different sensor positions. 3D point clouds extracted thanks to tomographic processing methods are often corrupted by noise and artifacts which need to be corrected. In this paper, we propose a simple convex optimization formulation that exploits the geometric constraint that the line of sight between a sensor and a surface measurement must be unobscured. We demonstrate the ability of our method to denoise point clouds and fill holes on both synthetic and experimental DLR E-SAR data.

Index Terms— TomoSAR, Free Space Constraint, Convex Optimization

1 Introduction

Tomographic SAR (TomoSAR) allows to reconstruct reflectivity profiles in elevation by using several images of a scene acquired with different viewing angles [1]. It allows to image volume structures such as forest canopy and separate scatterers in layover areas which makes it an interesting technique for the 3D imaging of urban areas. In this last case, 3D point clouds may be obtained by extracting peaks from the tomograms. For airborne data, artifacts may occur in these point clouds due to uncertainty on sensor position, phase ambiguities and multiple-bounce scattering effects. Moreover, decorrelation effects corrupt the interferometric phases and result in noisy positions of the extracted scatterers. Holes can also appear in the data due to very low coherence areas or when the estimated number of scatterers in wrong. For these reasons, it is important to develop methods which allow to filter or regularize the data without destroying the spatial content such as sharp discontinuities due to the objects present in the scene.

In this paper, we propose a framework to regularize the point clouds with a fast algorithm based on convex optimization. Our method is suitable for the reconstruction of surface-like scattering areas and allows to recover sharp height jumps and fill holes thanks to a regularization term based on nonlinear gradient penalization and the incorporation of confidence weights on the data points. Moreover, a projection of the points into a height map allows to improve the reconstruction by considering geometric visibility constraints.

This paper was supported by grants (HE 2459/21-1 and DH 75/2-1) from the Deutsche Forschungsgemeinschaft (DFG). The authors would like to thank DLR for providing the E-SAR dataset.

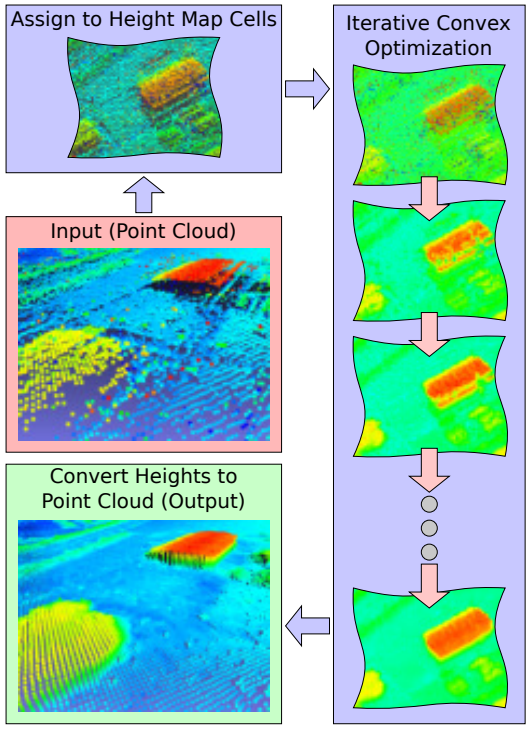
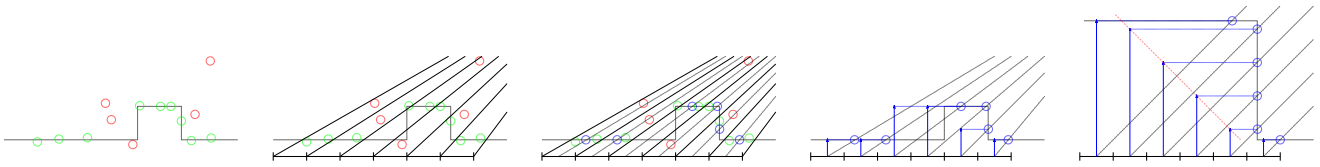


Fig. 1: Processing flow of our method from TomoSAR point cloud input to filtered output. The full point clouds, in 3D, can be inspected on our project website [2].

2 Prior Work

SAR tomography being a recent discipline, very few works consider the processing of the corresponding 3D point clouds. However, 3D reconstruction from InSAR data has been an active field for several years. An approach for the joint phase and intensity denoising of urban InSAR data considering a Markov Random Field model had been introduced in [3]. An extension to multi-channel InSAR data has been proposed in [4], which assumed independent channels. In [5] a new method allowed to exploit channel correlation. In [6], a bi-



(a) Building and ground plane as seen from a sideways looking SAR sensor resulting in correct (green circles) and noisy (red circles) point measurements.

(b) Association of the points with height map grid cells.

(c) Estimation of a new, regularized height and thus point (blue circles) for each height map cell. Only one height per cell and thus view ray is allowed.

(d) The height map values do not encode distances from the sensor, but heights over ground. Note that horizontal regions result in areas of homogeneous height.

(e) The gradients of slopes (e.g. vertical facades) in the height map must exceed a certain minimal negative gradient (red line) or result in overhangs.

Fig. 2: Surface representation

lateral filtering approach allowed to reduce the noise in TomoSAR 3D points by filtering the data in the covariance matrix domain. First approaches in denoising airborne tomographic SAR point clouds by primitive extraction have been proposed in [7] and extended to the extraction of buildings in [8]. In [9] a method to extract facades and building roofs from high resolution satellite images has been proposed.

3 Proposed Method

Our approach uses the primal dual optimization framework presented in [10]. Within this framework, an energy formulation composed of regularization and data terms is minimized. In the following, we will first describe our surface representation (see Sec. 3.1). Building on the properties of this representation we will describe our data and regularization terms (see Sec. 3.2) and then show how this energy fits into the primal dual optimization framework (see Sec. 3.3).

3.1 Surface Representation

In the following, we assume that our initial point cloud has been first projected from the slant range to the ground range geometry. We make the simplifying assumption, that the radar does not penetrate the structures that we are interested in (buildings and ground) and that each surface point is "seen" only from the point on the flight track closest to it (instead of being focused from an entire range). These assumptions imply that along each "view ray" from the sensor, only one distance measurement can be true. We enforce this by representing the surface as a special form of height map (see Fig. 2) which can be intuitively understood as an image of height values as seen from the sensor. Each pixel of this height map corresponds to a view ray from the sensor which for airborne SAR is realized through an orthographic azimuth projection and a perspective range projection.

The points of the input point cloud are assigned to their corresponding height map cells to constitute the data term in the following optimization. The heights in the height map are stored as actual vertical heights over ground and not as distances from the sensor. This has the nice property, that flat surfaces have equal height, even when "seen" at an angle.

3.2 Data and Regularization Terms

The choice of surface representation already prevents multiple depths per view ray making explicit enforcement of this constraint unnecessary. In addition, we formulate and subsequently minimize an energy over the height map values $x_{i,j}$

$$E(x) = \sum_{i,j} E_R(\operatorname{grad} x_{i,j}) + E_D(x_{i,j}). \tag{1}$$

The data term E_D is the weighted sum of the absolute distances between the estimated vertical height of each pixel $x_{i,j}$ and the heights of input points $p_{i,j,k}$ that fall into the pixel i, j, if any. The weights of the input points $w_{i,j,k}$ allow to assign a confidence to each estimated point. In this work, we use the pseudo-power of the peak which is a by-product of the MUSIC algorithm employed for tomographic focusing. We employ L_1-Norm instead of L_2-Norm to be more outlier resistant

$$E_D(x_{i,j}) = \sum_k w_{i,j,k} \left| x_{i,j} - p_{i,j,k} \right|. \tag{2}$$

The regularization term E_R is chosen such as to penalize deviations from horizontal surfaces and disallow overhangs. Both can be achieved by considering the gradients of the height map $\operatorname{grad} x_{i,j}$. The gradients are computed from the finite differences $\operatorname{grad}_a x_{i,j} = x_{i+1,j} - x_{i,j}$ and $\operatorname{grad}_r x_{i,j} = x_{i,j+1} - x_{i,j}$. Since azimuth and range direction in the height map have different geometric meanings (one corresponds to an orthographic projection, the other to a perspective projection), we consider the azimuth and range components of the gradients independently and weight them against each other with a weight factor β:

$$E_R(\operatorname{grad} x_{i,j}) = \beta \cdot |\operatorname{grad}_a x_{i,j}|_a + |\operatorname{grad}_r x_{i,j}|_r \tag{3}$$

In the range direction, we penalize non-zero gradients to encourage horizontal surfaces. To allow height jumps, e.g. for buildings and bridges, we use an L_1 norm on the gradient. Since overhanging structures shall be prevented, we penalize any gradient smaller than a certain minimal negative value α with infinite energy (compare Fig. 2e).

$$|x|_r = \begin{cases} |x| & \text{if } x \geq \alpha < 0 \\ \infty & \text{otherwise} \end{cases} \tag{4}$$

Algorithm 1 Adapted algorithm 1 from [10]

Input: $\tau, \sigma, \theta, \beta, \lambda$, coordinates of sensor track to compute α
Initialize: $\mathbf{x}_0 \Leftarrow$ highest data point of cell or average height of all points if cell empty
Initialize: $\bar{\mathbf{x}}_0 \Leftarrow \mathbf{x}_0$
Initialize: $\mathbf{y}_0 \Leftarrow \mathbf{0}$
for $n : [0..N]$ **do**
$\mathbf{y}_{n+1} \Leftarrow \text{prox}_{\sigma,F^*}(y_n + \sigma \underbrace{\mathbf{K}\bar{\mathbf{x}}_n}_{\text{grad}\,\bar{x}_n})$
$\mathbf{x}_{n+1} \Leftarrow \text{prox}_{\tau,G}(x_n - \tau \underbrace{\mathbf{K}^*\mathbf{y}_{n+1}}_{-\text{div}\,y_{n+1}})$
$\bar{\mathbf{x}}_{n+1} \Leftarrow \mathbf{x}_{n+1} + \theta\,(\mathbf{x}_{n+1} - \mathbf{x}_n)$
end for

The value of α is only locally constant and has to be computed for each pixel depending on its position in the range direction.

In the azimuth direction, we also penalize non-zero gradients. We experimented with the L_1 norm as well, but found that the Huber norm provided better results for angled building facades:

$$|x|_a = \begin{cases} \frac{1}{2} \cdot x^2 & \text{if } |x| < \lambda \\ \lambda\left(|x| - \frac{1}{2}\lambda\right) & \text{otherwise} \end{cases} \tag{5}$$

3.3 Optimization

The energy in Eq. 1 can be reformulated into a form that fits the optimization framework of [10]:

$$E(\mathbf{x}) = \underbrace{F(\mathbf{K}\mathbf{x})}_{\sum_{i,j} E_R(\text{grad}\,x_{i,j})} + \underbrace{G(\mathbf{x})}_{\sum_{i,j} E_D(x_{i,j})} \tag{6}$$

In this formulation, the computation of the finite differences of the gradients $\text{grad}\,x_{i,j}$ can be thought of as a multiplication of the concatenated height values $\mathbf{x}$ with a specifically crafted sparse matrix $\mathbf{K}$ (compare, e.g., with the examples in [10]). The function F then takes on the role of the regularization term E_R while the function G represents the data term E_D.

An in depth explanation of the optimization framework in [10] is beyond the scope of this paper and we refer interested readers to the original paper. However, in order to apply the optimization framework, as sketched in Alg. 1, solutions for the following three terms are needed: a) the adjunct $\mathbf{K}^*$ of $\mathbf{K}$, b) the resolvent operator prox_{σ,F^*}, and c) the resolvent operator $\text{prox}_{\tau,G}$. Here, as F^* we denote the convex conjugate of F. The resolvent operator for a scale ρ and a function $H(\mathbf{x})$ is defined as:

$$\text{prox}_{\rho,H}(\mathbf{x}) = \arg\min_{\widetilde{\mathbf{x}}} \left(\frac{|\widetilde{\mathbf{x}} - \mathbf{x}|^2}{2\rho} + H(\widetilde{\mathbf{x}}) \right) \tag{7}$$

Due to space limitations, we simply state the final formulas here and refer to our project website [2] for the full derivations. The adjunct $\mathbf{K}^*$ of $\mathbf{K}$ is the negative divergence operator ($(\mathbf{K}^*\mathbf{y})_{i,j} = -\text{div}\,y_{i,j}$). The resolvent operators can be computed pixel-wise and, for the regularization, can be split into its azimuth and range parts.

The pixel-wise azimuth part $\text{prox}_{\sigma,F_a^*}$ is the default resolvent operator for the Huber norm (see, e.g., [10]), except for the balancing factor β. The pixel-wise range part resolves to a simple per pixel case selection:

$$\text{prox}_{\sigma,F_r^*}(y_{i,j}) = \begin{cases} y_{i,j} & \text{if } |y_{i,j}| \le 1 \\ 1 & \text{if } y_{i,j} > 1 \\ -1 & \text{if } y_{i,j} < -1, y_{i,j} - \alpha\sigma \ge -1 \\ y_{i,j} - \alpha\sigma & \text{if } y_{i,j} < -1, y_{i,j} - \alpha\sigma < -1 \end{cases} \tag{8}$$

The pixel-wise resolvent operator for the data term $\text{prox}_{\tau,G}(x_{i,j})$ is slightly more complicated and has no closed form solution. We compute the minimum for each height interval between sorted data points $p_{i,j,k}$ and $p_{i,j,k+1}$ as well as for the height of each data point $p_{i,j,k}$ individually and then pick the lowest minimum. For details, see [2].

Despite the complexity of the involved math, the resulting algorithm is surprisingly simple. Since most operations are pixel wise, it is easy to exploit multi-core CPU or even GPU architectures.

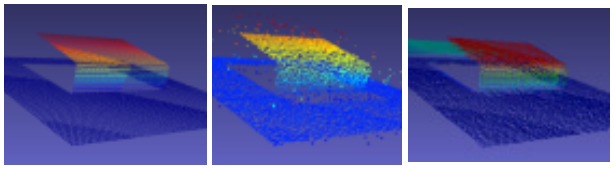
4 Experiments

4.1 Synthetic data

To evaluate the method we have simulated a scene composed of a simple 3D object representing a building. The 3D model has been discretized assuming rectilinear SAR trajectories and a stack of 6 images has been generated according to a layover model for distributed targets following the procedure established in [6]. In this work, we simulate baseline decorrelation, resulting in noisy heights. To assess the robustness of our algorithm against errors in the estimated number of scatterers, we have allowed randomness in the model order. To do so, 10% of pixels were randomly picked and their model order was replaced by a uniformly distributed random number. Next, the MUSIC algorithm has been applied to each pixel with the previously generated model order to retrieve the point cloud. Then, we have applied our reconstruction method. Fig. 3 shows the ground truth, simulated and processed point clouds. It can be noted that the contamination of the data results in wrongly estimated points besides the decorrelation noise. It may be observed that our method allows to successfully reconstruct the object and preserve the facade. Moreover, the method allows to retrieve a high point density in the ground part just in front of the facade. This part was almost empty in the simulated image due to erroneous tomographic estimations in the layover area. Furthermore, our algorithm shows robustness against a wrongly estimated number of scatterers, which is an important point when considering experimental data.

4.2 Experimental E-SAR data

We have applied the method to a point cloud obtained by tomographic processing of the experimentally acquired L-band



(a) Ground truth (b) Input points (c) Processed points

Fig. 3: Results on simulated data.

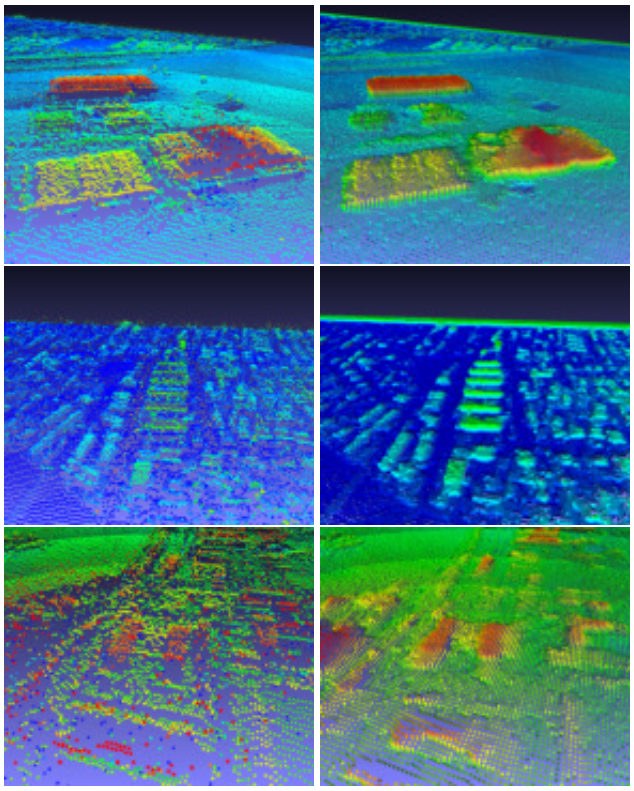


Fig. 4: Results on the point cloud of the E-SAR Oberpfaffenhofen dataset over 3 cropped areas. Original noisy points (left column) and points after applying our method (right column).

DLR E-SAR dataset of the Oberpfaffenhofen (1998) scene. The MUSIC method has been used for TomoSAR focusing, up to three scatterers. Fig. 4 compares the original noisy points with the ones after applying our new method. It may be observed that our method allows to smooth the flat areas while preserving the sharp height discontinuities between ground parts and building roofs. Moreover, thanks to the spatial regularization term and the confidence based weights, holes that appeared in the original data have been filled. This even holds for buildings facades which are facing the sensor, even if they are not appearing in the noisy input data.

5 Conclusion and Future Work

In this paper we have introduced a new method to regularize point clouds from TomoSAR data. The approach relies on an energy formulation which incorporates spatial regularization by imposing a gradient based prior on the reconstructed points. Moreover, by defining confidence weights on the estimated points, the data fidelity terms allows to fill holes in the point cloud and restore sharp height discontinuities. By enforcing visibility constraints, it allows the completion of facades even if few points are present and is therefore suitable for surface reconstruction in urban areas. Thanks to the use of a recent convex optimization framework, the algorithm is simple to implement, fast and highly parallelizable. The method has some limitations, namely its tendency to over-smooth not well supported vertical structures as the prior prefers horizontal, flat surfaces. The utilized projection from the sensor's "point of view" also makes extensions to multi-aspect scenarios non obvious.

In future works we will provide quantitative evaluation with more complex simulated scenes. We will also compare the method to other approaches.

6 References

[1] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, 2000.

[2] Andreas Ley, "Project website," `http://andreas-ley.com/projects/FreeSpaceSAR/`, 2017.

[3] L. Denis, F. Tupin, J. Darbon, and M. Sigelle, "Joint regularization of phase and amplitude of InSAR data: Application to 3-D reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 11, pp. 3774–3785, Nov 2009.

[4] A. Shabou, F. Baselice, and G. Ferraioli, "Urban digital elevation model reconstruction using very high resolution multichannel InSAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4748–4758, Nov 2012.

[5] F. Baselice, A. Budillon, G. Ferraioli, V. Pascazio, and G. Schirinzi, "Multibaseline SAR interferometry from complex data," *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, vol. 7, no. 7, pp. 2911–2918, July 2014.

[6] O. D'Hondt, S. Guillaso, C. Lopez-Martinez, and O. Hellwich, "Evaluation of a bilateral filtering approach for tomographic SAR denoising," in *EUSAR 2016*, June 2016, pp. 1–4.

[7] O. D'Hondt, S. Guillaso, and O. Hellwich, "Geometric primitive extraction for 3D reconstruction of urban areas from tomographic SAR data," in *JURSE 2013*, 2013, pp. 206–209.

[8] O. D'Hondt, S. Guillaso, and O. Hellwich, "Towards a semantic interpretation of urban areas with airborne synthetic aperture radar tomography," *ISPRS Annals*, vol. III-7, pp. 235–242, 2016.

[9] M. Shahzad and X. X. Zhu, "Automatic detection and reconstruction of 2-D/3-D building shapes from spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1292–1310, March 2016.

[10] Antonin Chambolle and Thomas Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *J. Math. Imaging Vision*, vol. 40, no. 1, pp. 120–145, 2011.