NON-LOCAL SAR TOMOGRAPHY FOR LARGE-SCALE URBAN MAPPING

Yilei Shi [1], Yuanyuan Wang [2], Xiao Xiang Zhu [2,3], Richard Bamler [1,3]

[1] Chair of Remote Sensing Technology (LMF), Technical University of Munich, Munich, Germany
[2] Signal Processing in Earth Observation (SiPEO), Technical University of Munich, Munich, Germany
[3] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Wessling, Germany

ABSTRACT

Multi-baseline synthetic aperture radar (SAR) interferometric techniques, such as SAR tomography, is well established for 3-D reconstruction in the urban area. These methods usually require fairly large interferometric stacks (> 20 images) for a reliable reconstruction. Hence, they are usually not directly applicable for large-scale 3-D urban mapping using TanDEM-X data where only a few acquisitions are available in average for each city. This work proposes a new SAR tomographic processing framework to those extremely small stacks. The applicability of the algorithm is demonstrated using a TanDEM-X multi-baseline stack with five bistatic interferograms over the whole city of Munich, Germany. Systematic comparison of our result with TanDEM-X raw digital elevation models (DEM) and airborne LiDAR data shows that the relative height accuracy is two meters, which outperforms the TanDEM-X raw DEM. The promising performance of the proposed algorithm paved the first step towards high quality large-scale 3-D urban mapping.

Index Terms— TomoSAR, Non-Local filtering, Robust estimator

1. INTRODUCTION

Synthetic aperture radar tomography (TomoSAR) is an advanced SAR interferemetric technique that is able to reconstruct the 3-D distribution of scatterers and retrieve the elevation profile orthogonal to the radar line of sight (LOS). Among the many multi-baseline InSAR techniques, TomoSAR is the only one that strictly reconstructs the full reflectivity along the third dimension elevation. SAR tomography and its differential form (D-TomoSAR) have been extensively developed in last two decades [1] [2] [4] [5] [6]. They achieve promising results on 3-D reconstruction of urban areas, especially when using high resolution data like TerraSAR-X [7] or COSMO-Skymed [8]. In last few years, new algorithms by taking advantage of recent developments in signal processing such as sparse reconstruction and compressive sensing (CS) [9] [10] can provide height estimates with unprecedented accuracy compared to the state-of-the-art multi-baseline InSAR algorithms and shows the super-

resolution (SR) power, which is very important for urban areas, since amount of areas have layover effects. Apart from the algorithmic development study was also conducted on TomoSAR using TanDEM-X bistatic data. It was shown that better height accuracy can be achieved with respect to a TerraSAR-X monostatic stack, because of the higher SNR and general quality of TanDEM-X bistatic data [11].

Although TanDEM-X bistatic data has many advantages, there is only a limited number of acquisitions available for most areas. For a reliable reconstruction, SAR tomography usually requires fairly large interferometric stacks, because the variance of the estimates is asymptotically related to the product of SNR and the number of acquisitions. Therefore, they are not directly applicable to InSAR stack with only a few acquisitions [12].

In this work, we follow the concept of non-local compressive sensing TomoSAR in [13] and propose a new framework of spaceborne multi-baseline SAR Tomography with very small TanDEM-X bistatic stacks, i.e. three to five interferograms. The framework includes non-local filtering, spectral estimation, model selection and robust height estimation. We choose Munich city as a test site and compare the TomoSAR point cloud generated by the proposed framework, TanDEM-X DEM product, and LiDAR data.

2. NON-LOCAL TOMOSAR

In this section, we introduce the non-local TomoSAR framework. It consists of several steps: (1) non-local filtering; (2) spectral estimation; (3) model selection; (4) robust height estimation.

2.1. SAR imaging model

The typical multi-baseline SAR imaging model can be expressed as follows:

$$g_n = \int_{\Delta s} \gamma(s) \cdot \exp(j2\pi\xi_n s)ds \tag{1}$$

where g_n is the complex-valued measurement of an azimuth-range pixel in the nth acquisition. $\gamma(s)$ represents the reflectivity function along elevation s with an extent of Δs. The

spatial frequency $\xi_n = 2b_n/\lambda r$ is proportional to the corresponding baseline b_n. λ is the wavelength of the radar signals and r denotes the range between radar and the observed object, respectively.

In the presence of noise ε, the discrete-TomoSAR system model can be rewritten as:

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \varepsilon \tag{2}$$

where $\mathbf{g}$ is the measurement vector with N elements, and $\boldsymbol{\gamma}$ is the reflectivity function along elevation uniformly sampled at $s_l(l = 1, 2, ..., L)$. $\mathbf{R}$ is an $N \times L$ irregularly sampled discrete Fourier transformation mapping matrix.

2.2. Non-Local Procedure

Since we have only limited number of acquisitions for large-scale area, the SNR need to be dramatically increased in order to obtain the required accuracy. As shown in [13], non-local procedure is efficient way to increase the SNR of interferograms without notable resolution distortion. The NL-means concept redefines the neighborhood of a pixel c in a very general sense as any set of pixels s in the image (local or non-local) such that a small patch around s is similar to the patch around c. It can combine similar patches into a weighted maximum likelihood estimator

$$\hat{\boldsymbol{\Theta}}_c = \arg\max \sum_s \mathbf{w}(i_s, j_s) \log p(\mathbf{g}_s | \boldsymbol{\Theta}) \tag{3}$$

Where weights $\mathbf{w}(i_s, j_s)$ depends on the statistical model of the imaging process [13]. $\mathcal{N}(.)$ is the non-local estimator and $\mathcal{N}(\mathbf{g}) = f(\hat{\boldsymbol{\Theta}})$. The expression $\hat{\boldsymbol{\Theta}} = (\hat{\psi}, \hat{\mu}, \hat{\sigma}^2)$ denotes the parameters, where $\hat{\psi}$ is the estimate of the interferometric phase, $\hat{\mu}$ is the coherence magnitude, and $\hat{\sigma}^2$ is variance.

2.3. Spectral Estimation

After the non-local procedure, spectral estimation is applied. The most relevant spectral estimation algorithms, including singular value decomposition (SVD), compressive sensing (CS) are introduced in the following.

- SVD:

$$\hat{\boldsymbol{\gamma}} = \left(\mathbf{R}^{\mathrm{H}}\mathbf{C}_{\epsilon\epsilon}^{-1}\mathbf{R} + \mathbf{C}_{\gamma\gamma}^{-1}\right)^{-1} \mathbf{R}^{\mathrm{H}}\mathbf{C}_{\epsilon\epsilon}^{-1}\mathcal{N}(\mathbf{g}) \tag{4}$$

- CS:

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}}\{\|\mathbf{R}\boldsymbol{\gamma} - \mathcal{N}(\mathbf{g})\|_2^2 + \lambda\|\boldsymbol{\gamma}\|_1\} \tag{5}$$

The choice of different combinations of spectral estimators depends on the required accuracy, the computational time and others. We follow the procedure proposed in [14]. It consists of three steps: first-order spectral estimation, single and

double scatterers discrimination, and higher order spectral estimation. It uses the well-established and computationally efficient first-order spectral estimator to obtain a prior knowledge of the estimates, followed by the linear method and CS-based algorithm applied to pre-classified different groups of pixels. This approach speeds up the processing by reducing the percentage of pixels that require sparse reconstruction. The sparse reconstruction can be further solved by the approach in [15].

2.4. Model Selection

The abovementioned spectral estimators retrieve a nonparametric reflectivity profile. Since our data is in urban area, we assume only a few dominant scatterers exist along the reflectivity profile. Therefore, we employ model order selection to determine the number of scatterers $\hat{K}$ as well as their elevation in one azimuth-range pixel [4]. The estimator can be expressed as follows.

$$\hat{K} = \arg\min_K \{-2\ln p(\mathbf{g}|\boldsymbol{\theta}) + 2C(K)\} \tag{6}$$

where $C(k)$ is a complexity penalty that trades off between how well the model fits the data and the complexity of the model.

2.5. Robust Height Estimation

To tackle the possible remaining outliers in the height estimates, the final height will be fused from the result of multiple neighbouring pixels as a post-processing. But instead of simple averaging, the height will be adjusted robustly using an *M-estimator*. Instead of minimizing the sum of squared residuals in averaging, M-estimator minimizes the sum of a customized function $\rho(.)$ of the residuals:

$$\tilde{s} = \arg\min_s \sum_i \rho(\hat{s}_i - s), \tag{7}$$

where $\hat{s}_i$ is the elevation estimates of the ith neighbouring pixel. It is shown that the close-formed solution of Eq. (7) is simply a weighted averaging of the heights of the neighbouring pixels. The weighting function can be expressed as follows, if the derivative of $\rho(x)$ exists.

$$w(x) = \frac{-\partial\rho(x)}{x\partial x} \tag{8}$$

The robust estimated height can be written as follows:

$$\tilde{h} = \frac{\sum_i w(x_i) \cdot \hat{h}}{\sum_i w(x_i)} \tag{9}$$

where $\hat{h} = \hat{s} \cdot \sin\theta$, and θ is incident angle.

3. PRACTICAL DEMONSTRATION

3.1. Data Description

We make use of a stack of five co-registered TanDEM-X bistatic interferograms to evaluate the proposed algorithm. The dataset is over Munich, Germany, with a slant range resolution of 1.2 m and an azimuth resolution of 3.3 m. The images were acquired from July 2016 to April 2017. The

r	λ	θ	Δb	N
698 km	3.1 cm	50.4°	187.18 m	5

Table 1. Parameters of Tandem-X Acquisition of Munich

most pertinent parameters of a TanDEM-X bistatic stripe map acquisition of Munich are listed in Table 1.

3.2. Visual Comparison with TanDEM-X raw DEM

In this work, the TanDEM-X raw DEM is adopted for visual comparison with TomoSAR point clouds of the test area, which is generated by two TanDEM-X bistatic acquisitions using the Integrated TanDEM-X Processor (ITP).
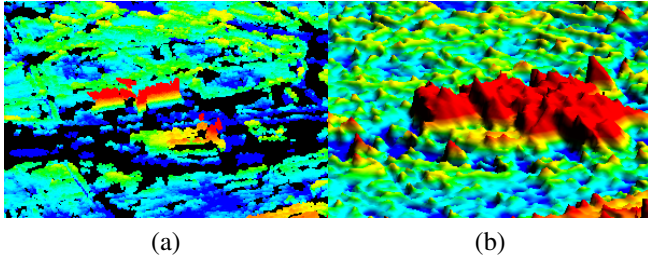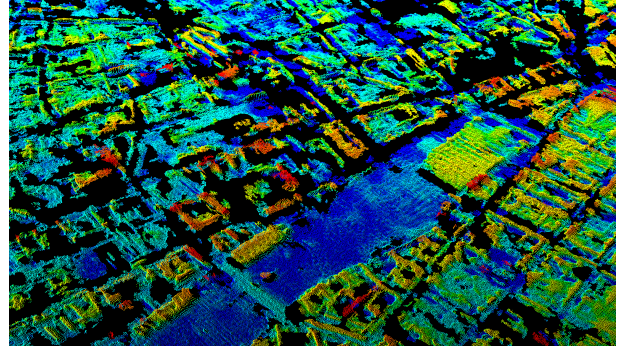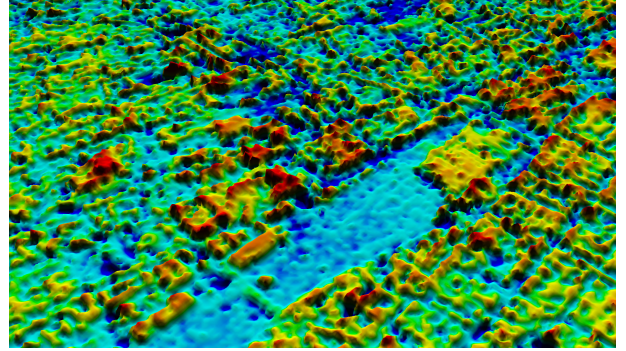

| (a) | (b) |

Fig. 1. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM, close-up 3-D view over the area of European bureau of patent. (a) TomoSAR point clouds. (b) TanDEM-X DEM.

Fig. 1 shows the area of European bureau of patent. As one can see in Fig. 1 (b), due to the complex structure of the building and low resolution, TanDEM-X DEM merge several buildings together and exhibits large error on the height of the buildings.

Fig. 2 shows the visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM with a close-up 3-D view over the area of Munich central station. It is clear that NL-TomoSAR result can show more detailed structures, such as the bridge, the central station, and roads. There are several reasons for the blur in case of TanDEM-X DEM: layover of building superimposes signal from roads and shadow cast behind buildings lead to many noisy areas in cities.

(a)

(b)

Fig. 2. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM, close-up 3-D view over the area of Munich central station. (a) TomoSAR point clouds. (b) TanDEM-X DEM.

3.3. Quantitative Comparison of Individual Structure

In order to evaluate the estimation accuracy, nine test sites with high average SNR have been chosen for individual quantitative comparison. The summary of the results is shown in Tab. 2.

	S1	S2	S3	S4	S5	S6	S7	S8	S9
T	0.69	0.75	0.90	0.67	0.96	0.67	0.60	0.94	0.89
D	6.02	4.21	6.62	2.17	2.02	7.99	8.46	5.41	2.88

Table 2. Statistics of quantitative comparison of nine test structures. Relative height differences [m] compared with reference (LiDAR). T (TomoSar), D (DEM).

From Tab. 2 we can see that the height differences between TomoSAR result and LiDAR data are within one meter and the height differences between TanDEM-X DEM product and LiDAR data vary from 2.5 m to 8.5 m.

3.4. Average accuracy

In order to have an assessment of the overall accuracy in a city scale, we compared all the 36,499 buildings in the area with the LiDAR point cloud. 38.7% buildings are within 1 m accuracy. 62.8% are within 2 m accuracy. However, the two datasets (TanDEM-X CoSSC and LiDAR) were acquired at different time. It is almost certain that changes happened during the period. Therefore, in order to obtain a more realistic assessment, we truncated the distribution of height difference at ±15 m. 34,054 buildings remains after the truncation. Their overall standard deviation is 1.96 m.

4. CONCLUSION

In this work, we propose a novel framework for TomoSAR with minimum number of acquisitions to obtain fast and accurate estimation of elevation without any priori knowledge and demonstrated the applicability of the algorithm with test site in Munich. Using the real data, it was shown that the proposed method outperform the state-of-the-art.

5. ACKNOWLEDGMENT

6. REFERENCES

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

[2] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 1, pp. 37-44, Jan. 2005.

[3] G. Fornaro, F. Lombardini, and F. Serafino, "Three-dimensional multipass SAR focusing: experiments with long-term spaceborne data," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 4, pp. 702-714, Apr. 2005.

[4] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296-4308, Dec. 2010.

[5] N. Ge, F. Rodriguez Gonzalez, Y. Wang, Y. Shi, and X. X. Zhu, "Spaceborne Staring Spotlight SAR Tomography– A First Demonstration With TerraSAR-X," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 11, no. 10, pp. 702-714, Oct. 2018.

[6] X. X. Zhu, Y. Wang, S. Montazeri, N. Ge, "A Review of Ten-Year Advances of Multi-Baseline SAR Interferometry Using TerraSAR-X Data," *Remote Sensing*, vol. 10, no. 9, pp. 1374, Aug. 2018.

[7] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150-3157, Aug. 2012.

[8] G. Fornaro, A. Pauciullo, D. Reale, and S. Verde. "Multilook SAR Tomography for 3-D Reconstruction and Monitoring of Single Structures Applied to COSMO-SKYMED Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 7, pp. 2776-2785, Jul. 2014.

[9] X. X. Zhu and R. Bamler, "Tomographic SAR inversion by L_1 norm regularization – The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839-3846, Oct. 2010.

[10] A. Budillon, A. Evangelista, and G. Schirinzi, "Three-dimensional SAR focusing from multipass signals using compressive sampling," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 488-499, Jan. 2011.

[11] X. X. Zhu and R. Bamler, "Sparse tomographic SAR reconstruction from mixed TerraSAR-X/TanDEM-X data stacks," *IGARSS*, 2012.

[12] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247-258, Jan. 2012.

[13] Y. Shi, X. X. Zhu and R. Bamler, "Non-Local Compressive Sensing Based SAR Tomography," *IEEE Trans. Geosci. Remote Sens.*, 2018, in press.

[14] Y. Wang, X. X. Zhu, and R. Bamler, "An Efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1250-1254, Jul. 2014.

[15] Y. Shi, X.X. Zhu, W. Yin, and R. Bamler, "A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6148-6158, Oct. 2018.