Large-Scale Urban Mapping using Small Stack Multi-baseline TanDEM-X Interferograms

Yilei Shi [1], Yuanyuan Wang [2], Xiao Xiang Zhu [2,3], Richard Bamler [1,3]
[1] Chair of Remote Sensing Technology (LMF), Technische Universität München (TUM)
[2] Signal Processing in Earth Observation (SIPEO), Technische Universität München (TUM)
[3] Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR)
Munich, Germany (yilei.shi@tum.de)

***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. They are not directly applicable to SAR interferometric (InSAR) stack with only a few acquisitions, as the extremely small number of acquisitions can severely bias the estimates from the spectral estimators, such as beamforming which is often only asymptotically optimal. In addition, the number of images also causes severe ambiguity issue of the pixel with low signal-to-noise ratio. In this work, we propose a new processing framework of 3-D reconstruction with TomoSAR using extremely small stacks. Moreover, the applicability of the algorithm is demonstrated by exploiting TanDEM-X co-registered phase preserving single look slant range complex SAR images (CoSSC) over a large-scale test site of the whole Munich city, Germany. The reconstructed results have been systematically compared with global production of TanDEM-X digital elevation models (DEM) and LiDAR dataset, which show the potential of high quality large-scale 3-D urban mapping.**

***Index Terms*—SAR interferometric (InSAR), SAR tomography (TomoSAR), TanDEM-X, digital elevation models (DEM), 3-D urban mapping**

ALONG with the launch of TanDEM-X in 2010, for the first time there is a real multi-antenna system in space. It enables us to acquire data pairs simultaneously result in single-pass interferograms which have a high data quality and are free from deformation, atmosphere and temporal decorrelation. Interferometric data acquisition with the TanDEM-X satellite formation can be achieved in three different operational modes: Bistatic, Monostatic, and Alternating Bistatic Mode. Bistatic interferometry (Bistatic Mode) is characterized by the illumination of a scene by one transmitter and the simultaneous measurement of the same scene with two receivers [1]. In this way, it is possible to use single-pass spaceborne data covering a relatively large area for tomographic processing.

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). The repeat-pass multi-baseline SAR tomography has been intensively developed in last few years [2] [3] [4] [5] [7] [6] [8] and shows promising results on 3-D reconstruction of urban area [9]. The use of bistatic InSAR data introduces high-quality interferograms in the tomographic processing, yielding better height accuracy with respect to a monostatic stack. This fact has been already demonstrated using TanDEM-X single-pass image pairs for tomography [10].

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 ($>$ 20 images). Therefore, they are not directly applicable to InSAR stack with only a few acquisitions, as the extremely small number of acquisitions can severely bias the estimates from the spectral estimators. As shown in [11], it is asymptotically only the product of the number of acquisitions and SNR that determines the reconstruction quality, therefore the number of acquisitions and the signal-to-noise ratio (SNR) can be traded off against each other. Recent work [14] shows that SNR can dramatically increase by integrating non-local estimation into the inversion of TomoSAR and exhibits a reasonable reconstruction of buildings from only seven interferograms is feasible. In this work, we follow the concept of non-local compressive sensing TomoSAR in [14] and propose a new framework of spaceborne single-pass multi-baseline SAR Tomography with very small stacks, i.e. three to five interferograms.

The paper is organized as follows: In section II, the non-local TomoSAR framework is introduced; In section III, the estimation accuracy of TomoSAR with small stacks has been systematically studied; The experiments using real data, is presented in section IV; In section V, the quantitative validation is carried out; Finally, conclusions are given in section V.

I. Non-Local SAR Tomography

A. SAR Imaging

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 at an azimuth-range pixel for the nth acquisition at time $t_n (n = 1, 2, ..., N)$. $\gamma(s)$ 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 respective aperture position (baseline) b_n, where λ is the wavelength of the radar signals and r denotes the range between radar and the observed object, respectively.

978-1-7281-0009-8/19/$31.00 ©2019 IEEE

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.

B. Non-Local TomoSAR Framework

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 [14], 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 (WMLE)

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

The measure of the patch similarity that leads to the weights $\mathbf{w}(i_s, j_s)$ depends on the statistical model of the imaging process. $\mathcal{N}(.)$ 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.

After the non-local procedure, spectral estimation is performed.

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

The choice of different combinations of spectral estimators depends on the required accuracy, the computational time and others. Even for the compressive sensing based TomoSAR invension [12], we have proposed a fast and accurate algorithm to solve it [13].

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 [6]. The estimator can b expressed as follows.

$$\hat{K} = \arg\min_K \left\{-2 \ln p\left(\mathbf{g} | \boldsymbol{\theta}\right) + 2C(K)\right\} \tag{5}$$

II. Estimation accuracy of TomoSAR with small stacks

The estimation accuracy of TomoSAR has been systematically investigated in [11]. It shows that the elevation estimation accuracy and SR power depend asymptotically on the product $N \cdot$SNR. In this section, we analyze the estimation accuracy of TomoSAR with the extremely small number of interferograms, which is 3 to 5.

A. CRLB

The Cramer-Rao lower bound (CRLB) on elevation estimates of single scatterer can be shown to be [6]

$$\sigma_s = \frac{\lambda r}{4\pi \cdot \sigma_b \cdot \sqrt{2 \cdot \mathrm{SNR} \cdot N}} \tag{6}$$

where N is the number of acquisitions, SNR is the signal-to-noise ratio, and σ_b is the standard deviation of the baseline distribution.

For the double scatterers' case, the CRLB can be written as:

$$\sigma_{s_q} = c_0 \cdot \sigma_{s_q,0} \tag{7}$$

where $\sigma_{s_q,0}$ is the CRLB of the elevation estimates of the qth scatterer in the absence of the other one, and c_0 is the essential interference correction factor for closely spaced scatterers [11].

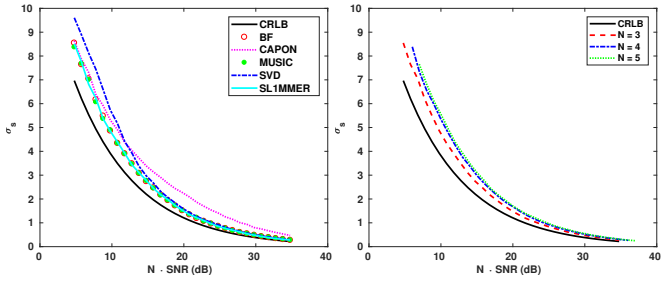
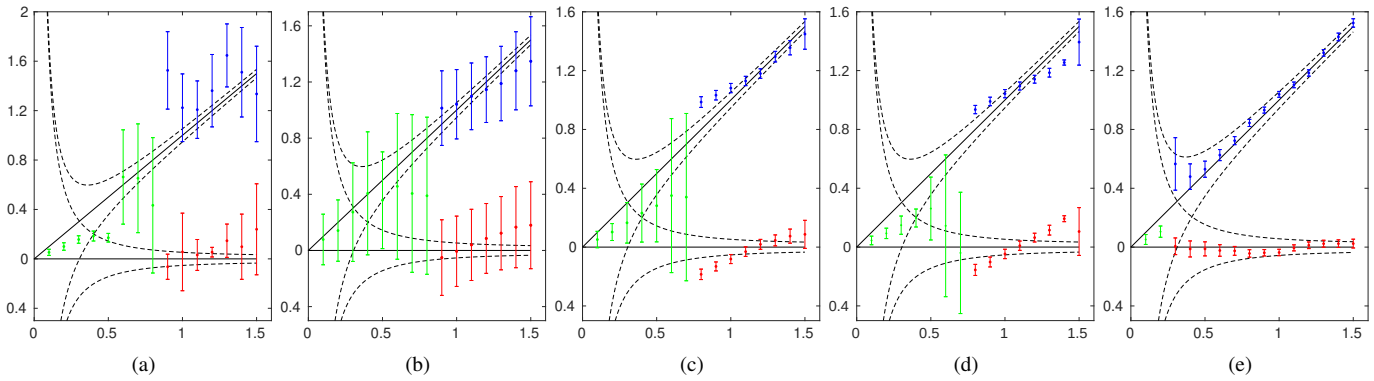


Fig. 1. Monte Carlo simulations of single scatterer with SNR in [0 30] (dB). (a) Comparison of CRLB with different spectral estimators with three acquisitions. (b) Comparison of CRLB using beamforming method with three to five acquisitions.

For the first test case, only one scatterer is placed at $s = 0$, and the SNR is in the range between 0 and 30 dB. A Monte Carlo simulation with 50,000 realizations per SNR value was performed to evaluate the CRLB of different schemes. Fig. 1 (a) shows a performance comparison between BF, CAPON, MUSIC, SVD, and CS on simulated data with three acquisitions for a single scatterer. As one can see, CAPON has worst estimation accuracy for the single scatterer. Since the number of acquisition is extremely small, the covariance matrix $\mathbf{C_{gg}}$ is therefore badly estimated. SVD has worse performance than other methods when SNR is smaller than 10 dB, (i.e., $N \cdot \mathrm{SNR} = 15$ dB). In contrast, BF, MUSIC, and SL1MMER have similar estimation accuracy. When $N \cdot$SNR is small, the estimation accuracy of these three methods has small deviation from CRLB. And when $N \cdot$SNR increases, they are going to collapse with CRLB. Fig. 1 (b) presents the estimation accuracy of beamforming with $N = 3, 4, 5$. If we choose three points on each curve with the same $N \cdot$SNR, it can be seen that the point on the curve with $N = 3$ has the smallest value of σ_s. According to Eq. (6), three points suppose to have the same value of σ_s. This evidence indicates that SNR has more effect on estimation accuracy when N is very small.

Then, we assume the situation with double scatterers inside an azimuth-range pixel: one scatterer located at the building



(a) (b) (c) (d) (e)

Fig. 2. Monte Carlo simulations of double scatterer with different normalized distances: $\kappa \in [0.1, 1.5]$.

facade and another from the ground with different normalized distances : $\kappa \in [0.1, 1.5]$; and a number of acquisitions, $N = 3 - 5$.

Fig. 2 shows the Monte Carlo simulations of double scatterers with different normalized distances : $\kappa \in [0.1, 1.5]$. The x-axis represents normalized true distance κ of simulated facade and ground. Y-axis is normalized estimated distance $\hat{\kappa}$ of simulated facade and ground. In each subplot, the two solid line segments mark the true elevation for facade and ground, respectively, while the dashed lines denote the true elevation plus and minus the CRLB, which is the same as in Fig. 2 (a). The blue dot marker denotes the estimated location of the facade and the error bar indicates the standard deviation of the estimates, whereas the red dot marker represents the estimated location of the ground. The green dot suggests that the detection rate of double scatterers is below 5% and denotes the estimated result of the single scatterer. Fig. 2 (b)-(f) show the estimated results by RELAX, CAPON, MUSIC, SVD and CS, respectively.

As one can see in Fig. 2 (b) (c), the results of RELAX and CAPON have large bias and standard deviation. In contrast, MUSIC exhibits good performance and the result has both small bias and standard deviation. Note that MUSIC algorithm is a parametric approach, which means the number of scatterers K should be given as a prior. The result of SVD is slightly worse than MUSIC, but it is a non-parametric approach and is no need to know the number of scatterers in advance. Nothing surprising, comparing to other methods, CS can give the best result, not only the accuracy of the estimation but also the super-resolution power.

III. Practical Demonstration and Validation

A. Data Description

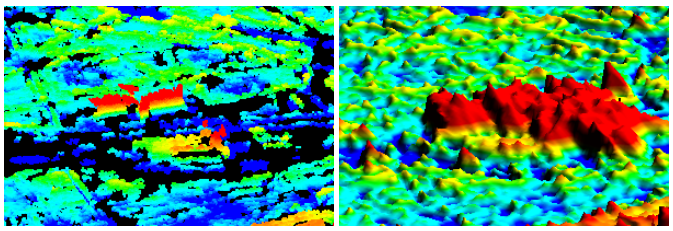
We make use of a stack of five co-registered TanDEM-X bistatic interferograms to evaluate the proposed algorithm. The dataset is over Munich, Germany, with a slant range resolution of 1.2 m and an azimuth resolution of 3.3 m. The images were acquired from July 2016 to April 2017. The most pertinent parameters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

TABLE I
Parameters of TanDEM-X Stripe Map Acquisition of Munich

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

B. Visual Comparison with TanDEM-X DEM

In this work, TanDEM-X raw DEM is adopted for visual comparison with TomoSAR point clouds of the test area.



(a) (b)

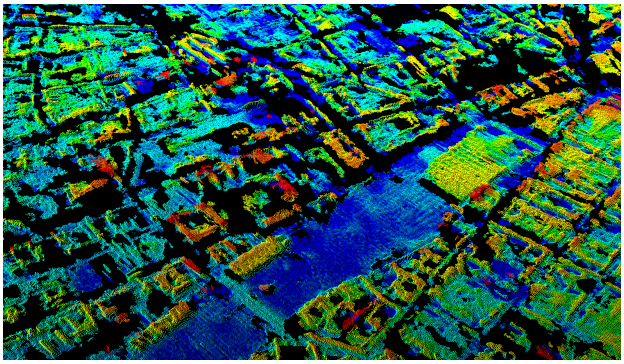
Fig. 3. Visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM, close-up 3-D view over the area of European bureau of patent. (a) TomoSAR point clouds. (b) TanDEM-X DEM.

Fig. 3 shows the area of European bureau of patent. As one can see in Fig. 3 (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.

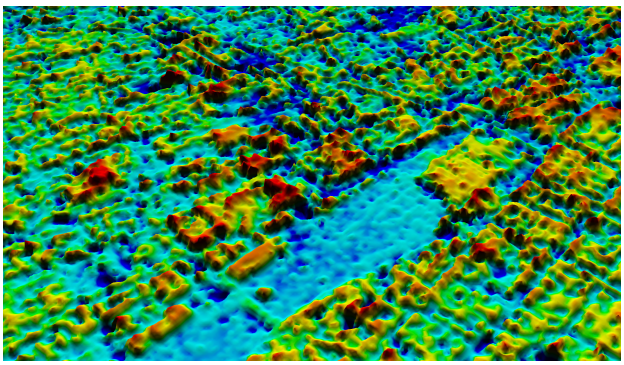
Fig. 4 shows the visual comparison of NL-TomoSAR point clouds and TanDEM-X DEM with a close-up 3-D view over the area of Munich central station. It is clear that NL-TomoSAR result can show more detailed structures, such as the bridge, the central station, and roads. There are several reasons for the 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.

C. Quantitative Validation

In this section, we have quantitatively compared the TomoSAR point clouds with TanDEM-X DEM and LiDAR point



(a)



(b)

Fig. 4. 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.

clouds. A LiDAR point cloud in Munich with centimeter accuracy was used as the reference point cloud. As different data sources have different coordinates and quality, we apply the following steps on the data. (1) Geocoding of TomoSAR point cloud; (2) Co-registration of different point clouds; (3) Object-based raster data generation; (4) Robust height estimation.

TABLE II
STATISTICS OF QUANTITATIVE COMPARISON OF NINE TEST STRUCTURES. RELATIVE HEIGHT DIFFERENCES [M] COMPARED WITH REFERENCE (LIDAR). T (TOMOSAR), D (DEM).

	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

From Tab. II 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. For the whole city, 34,054 buildings have been evaluated. The result shows that the standard deviation of height difference is 1.96 m.

IV. CONCLUSION

A new SAR tomographic inversion framework tailored for very limited number of measurements is proposed in this work. Our systematic investigation of the estimation accuracy of TomoSAR with a very small stack using simulated data. Experiments using TanDEM-X bistatic data shows the an relative height accuracy of 2 m can be achieved in large scale. Thus it demonstrate the proposed framework being a promising solution for high quality large-scale 3-D urban mapping.

ACKNOWLEDGMENT

This work is supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (no. ERC-2016-StG-714087, acronym: So2Sat), the Helmholtz Association under the framework of the Young Investigators Group "SiPEO" (VH-NG-1018, www.sipeo.bgu.tum.de), Munich Aerospace e.V. – Fakultät für Luft- und Raumfahrt, and the Bavaria California Technology Center (Project: Large-Scale Problems in Earth Observation). The authors thank Landesamt für Digitalisierung, Breitband und Vermessung Bayern provide the datasets.

REFERENCES

[1] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink, "TanDEM-X: A satellite formation for high-resolution SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 3317-3341, Oct. 2007.

[2] 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.

[3] F. Gini, F. Lombardini, and M. Montanari, "Layover solution in multibaseline SAR interferometry," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 4, pp. 1344-1356, 2002.

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

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

[12] 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.

[13] 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.

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