Nonlocal Compressive Sensing-Based SAR Tomography

Yilei Shi, *Member, IEEE*, Xiao Xiang Zhu , *Senior Member, IEEE*, and Richard Bamler , *Fellow, IEEE*

Abstract—Tomographic synthetic aperture radar (TomoSAR) inversion of urban areas is an inherently sparse reconstruction problem and, hence, can be solved using compressive sensing (CS) algorithms. This paper proposes solutions for two notorious problems in this field. First, TomoSAR requires a high number of data sets, which makes the technique expensive. However, it can be shown that the number of acquisitions and the signal-to-noise ratio (SNR) can be traded off against each other, because it is asymptotically only the product of the number of acquisitions and SNR that determines the reconstruction quality. We propose to increase SNR by integrating nonlocal (NL) estimation into the inversion and show that a reasonable reconstruction of buildings from only seven interferograms is feasible. Second, CS-based inversion is computationally expensive and therefore, barely suitable for large-scale applications. We introduce a new fast and accurate algorithm for solving the NL L1-L2-minimization problem, central to CS-based reconstruction algorithms. The applicability of the algorithm is demonstrated using simulated data and TerraSAR-X high-resolution spotlight images over an area in Munich, Germany.

Index Terms—Compressive sensing (CS), interferometric synthetic aperture radar (InSAR), nonlocal (NL) filtering, tomographic SAR (TomoSAR).

I. INTRODUCTION

SYNTHETIC aperture radar tomography (TomoSAR) is an advanced SAR interferometric technique that can not only retrieve 3-D spatial information but also assess the 4-D temporal information, e.g., deformation, in millimeter

scale, of individual buildings from meter-resolution SAR satellite data. Repeat-pass multibaseline TomoSAR has been intensively developed in the past two decades [1]–[6] and shows promising results on 3-D reconstruction of urban areas. However, for urban monitoring, there are still several issues that need to be solved: improving the elevation resolution, i.e., providing super-resolution (SR) for layover separation; achieving high 3-D localization accuracy even in the presence of unmodeled, non-Gaussian noise; and retrieving nonlinear motion, e.g., due to seasonal thermal dilation. Driven by these requirements, new algorithms have been invented in the past few years that take advantage of recent developments in signal processing [7]–[13], such as sparse reconstruction and compressive sensing (CS), and can provide height estimates with unprecedented accuracy compared with the state-of-the-art multibaseline InSAR algorithms [14].

However, CS-based TomoSAR still suffers from two problems for practical use. First, a large number of images are required, typically a stack of 20–100 images over the illuminated area. For instance, it is demonstrated in [15] that by using even the most efficient algorithms, such as nonlinear least squares and Scale-down by L_1 norm Minimization, Model selection, and Estimation Reconstruction (SL1MMER), a minimum number of 11 acquisitions are required to achieve a reasonable reconstruction in the interesting parameter range of spaceborne SAR. In [16], a joint sparsity concept was applied to obtain precise TomoSAR reconstruction with only six images by incorporating building *a priori* knowledge to the estimation. However, due to its demand on precise geometric prior, this method can only be used to reconstruct buildings where the geographic information system data are available. The second practical drawback of CS-based TomoSAR is that it suffers from a high computational expense and is hard to extend to large-scale practice. Wang *et al.* [17] proposed an efficient approach to address this issue, which uses the well-established and computationally efficient persistent scatterer interferometry to obtain *a priori* knowledge of the estimates, followed by the linear method and the CS-based SL1MMER algorithm applied to different preclassified groups of pixels. This approach speeds up the processing, but only to the extent that it reduces the percentage of pixels that require sparse reconstruction.

In this paper, we propose a novel framework for TomoSAR with a minimum number of acquisitions to obtain a fast and accurate estimation of elevation without any *a priori* knowledge. It is mainly motivated by the recent

advances in nonlocal (NL) means approaches [18], [19] in image restoration. NL-means approaches successfully achieve the state-of-the-art performance in image restoration [19] by seeking the correlation of image patches. As a common prior in natural images, the patch correlation should help increase the signal-to-noise ratio (SNR) of the original signal. As comprehensively investigated in [15], the product of the number of acquisitions and SNR determines the reconstruction quality, which means that an increase of SNR can dramatically reduce the number of acquisitions needed for reconstruction.

The main contributions of this paper are listed as follows.

1) We propose a novel framework, "NLCS-TomoSAR," to produce accurate estimates of scatterers' position without any *a priori* knowledge, using as few images as possible.

2) We further propose an efficient algorithm to solve the NLCS model, containing an optimized parallelization scheme for an NL process and a fast and accurate solver for complex-valued L_1 least squares minimization.

3) Systematic performance evaluation of the proposed approach has been carried out using both simulated and real data. The results show that the proposed method can produce very accurate estimations of elevation without notable resolution distortion.

This paper is organized as follows. In Section II, the SAR imaging model and TomoSAR inversion with CS approach are introduced. In Section III, a novel approach called "NLCS-TomoSAR" is introduced. Experiments using simulated data and real data are presented in Sections IV and VI. Finally, conclusions are given in Section VII.

II. CS-BASED SAR TOMOGRAPHY

A. SAR Imaging Model

The typical multibaseline SAR imaging model can be expressed as follows:

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

where g_n is the complex-valued measurement at an azimuth-range pixel for the nth acquisition at time $t_n(n = 1, 2, \ldots, N)$. The term $\gamma(s)$ 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, λ is the wavelength of the radar signals, and r denotes the range between radar and the observed object, respectively (see Fig. 1).

In the presence of noise $\boldsymbol{\varepsilon}$, the discrete-TomoSAR system model can be rewritten as

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \qquad (2)$$

where $\mathbf{g}$ is the measurement vector with N elements and $\boldsymbol{\gamma}$ is the reflectivity function along elevation uniformly sampled at $s_l(l = 1, 2, \ldots, L)$. $\mathbf{R}$ is an $N \times L$ irregularly sampled discrete Fourier transformation mapping matrix. The inherent (Rayleigh) elevation resolution ρ_s of the tomographic arrangement is related to the elevation aperture extent Δb
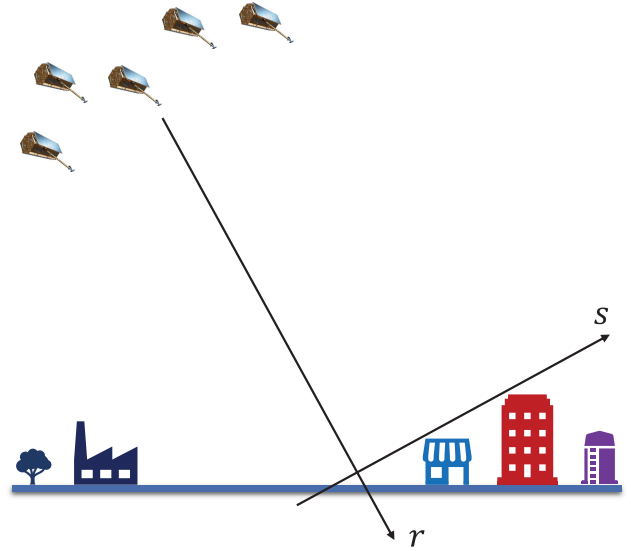
$$\rho_s = \frac{\lambda r}{2\Delta b}. \qquad (3)$$

Fig. 1.   TomoSAR imaging geometry with an artistic view of TerraSAR-X/TanDEM-X.

B. SL1MMER Algorithm

To solve (2), Zhu and Bamler [10] proposed SL1MMER. They demonstrated its SR power and robustness for spaceborne TomoSAR in [14]. The SL1MMER algorithm improves the CS algorithm and estimates these parameters in a highly accurate and robust way. It consists of three main steps.

1) L1LS minimization

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}}\{\|\mathbf{R}\boldsymbol{\gamma} - \mathbf{g}\|_2^2 + \lambda_1\|\boldsymbol{\gamma}\|_1\}. \qquad (4)$$

Model-order

2) selection

$$\hat{K} = \arg\min_{K}\{-2\ln p(\mathbf{g}|\boldsymbol{\theta}) + 2C(K)\}. \qquad (5)$$

3) Dealiasing

$$\hat{\boldsymbol{\gamma}} = (\mathbf{R}^H\mathbf{R})^{-1}\mathbf{R}^H\mathbf{g}. \qquad (6)$$

Here, K is the number of scatters. $C(k)$ is a complexity penalty, from which we can see that model selection is actually a tradeoff between how well the model fits the data and the complexity of the model. $p(\mathbf{g}|\boldsymbol{\theta})$ is the likelihood function, which is defined in Section II-C.

Within the framework of SL1MMER, sparse reconstruction and ordinary least squares join forces to incorporate both robust identification of scatterers' elevation positions and accurate amplitude estimation. However, as mentioned earlier, the CS-based approach has two downsides that can prevent its application. In order to solve those two issues, we need to analyze the estimation accuracy of TomoSAR.

C. Estimation Accuracy

Assume that $\boldsymbol{\theta}$ is a set of parameters for a given observation $\mathbf{g}$ and $p(\mathbf{g}|\boldsymbol{\theta})$ is the likelihood function. The Cramer–Rao lower bound (CRLB) $\mathbf{B}_{CR}$ can be calculated from the inverse of the Fisher information matrix $\mathbf{J}$, which is

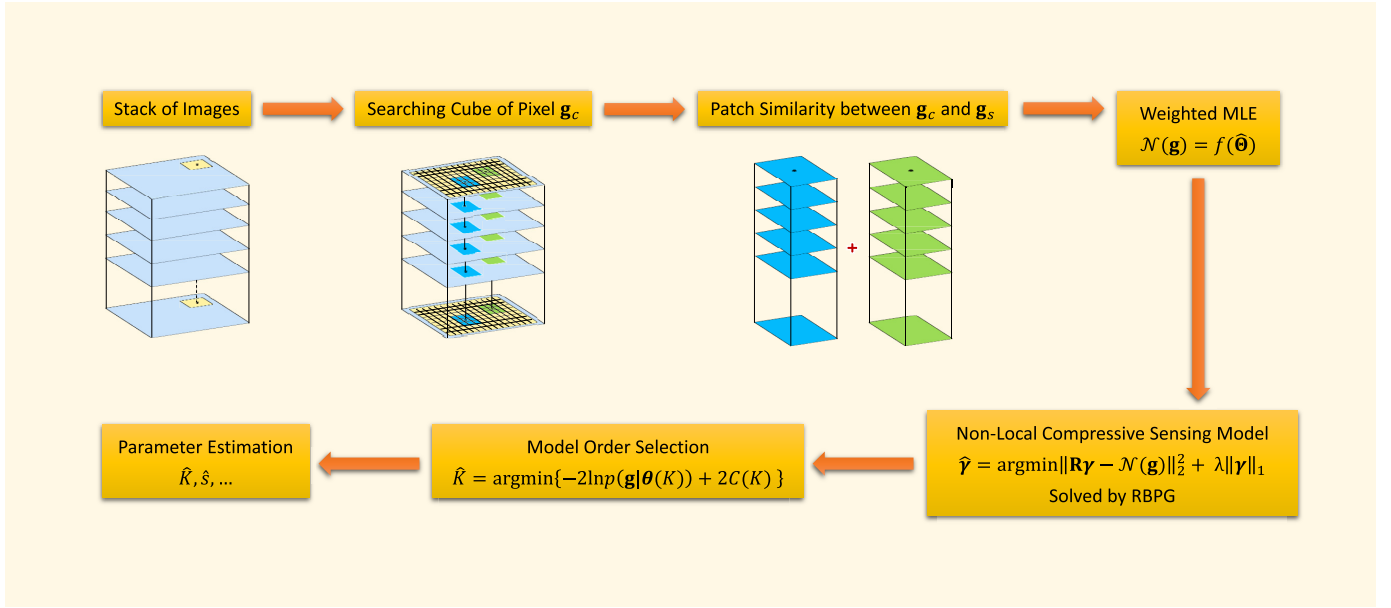$$\mathbf{B}_{CR} = \mathbf{J}^{-1} \qquad (7)$$

Fig. 2. Workflow of NLCS-TomoSAR.

and

$$\mathbf{J} = -E\left\{\frac{\partial^2 \ln p(\mathbf{g}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^H}\right\}. \tag{8}$$

Since the analytical inversion of $\mathbf{J}$ leads to a complicated expression, the relevant elements of the CRLB matrix can be retrieved by solving the inversion numerically. The CRLB on the elevation estimates of two scatterers can, therefore, be defined as

$$\sigma_{s_q} = c_0 \cdot \sigma_{s_q,0} \tag{9}$$

where

$$\sigma_{s_q,0} = \frac{\lambda r}{4\pi \cdot \sigma_b \cdot \sqrt{2 \cdot N \cdot \text{SNR}}} \tag{10}$$

is the CRLB of the elevation estimates of the qth scatterer in the absence of the other one and σ_b is the standard deviation of baseline. The essential interference correction factor for closely spaced scatterers is denoted by c_0. It has been systematically investigated in [15] that c_0 is almost independent of N and SNR, which is defined as

$$c_0 = \max\left\{\sqrt{\frac{40\kappa^{-2}(1-\kappa/3)}{9-6(3-2\kappa)\cos(2\Delta\varphi)+(3-2\kappa)^2}}, 1\right\} \tag{11}$$

where $\kappa = \Delta s/\rho_s$ is the normalized distance between two scatterers and $\Delta\varphi$ is the phase difference.

As shown in (10), the estimation accuracy of SL1MMER asymptotically depends on the product $N \cdot \text{SNR}$. Therefore, in order to maintain the estimation accuracy and reduce the number of measurements, SNR needs to be improved. A successful approach to reducing the noise as well as increasing the SNR is the NL framework, where the value is sum weighted with respect to the similarity between the central and other pixels in the search window. NL-means filtering is

consistent with the state of the art in image denoising and other applications. Hence, we introduce the NL framework into SL1MMER to achieve good performance with a minimal number of acquisitions.

III. NONLOCAL CS-BASED SAR TOMOGRAPHY

The NL concept proposed in [18] takes advantage of the high degree of redundancy of any natural image. This means that every feature edge, point, and so on in an image can be found similarly many times in the same image. Inspired by the neighborhood filters, such as boxcar and adaptive filters, 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 NL) such that a small patch around s is similar to the patch around c. Fig. 2 shows the workflow of the proposed NLCS-TomoSAR method.

A. Nonlocal Compressive Sensing

In the cases where there is no prior knowledge about the number of scatters and in the presence of measurement noise, the NLCS-based TomoSAR inversion can be written as

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}}\{\|\mathbf{R}\boldsymbol{\gamma} - \mathcal{N}(\mathbf{g})\|_2^2 + \lambda_1\|\boldsymbol{\gamma}\|_1\} \tag{12}$$

where $\mathcal{N}(.)$ is the NL 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 the variance, which will be introduced later. NL-means can combine similar patches into a weighted maximum likelihood estimator (WMLE)

$$\hat{\boldsymbol{\Theta}}_c = \arg\max \sum_s \mathbf{w}(i_s, j_s) \log p(\mathbf{g}_s|\boldsymbol{\Theta}). \tag{13}$$

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. In our case, it is derived from the InSAR statistics.

B. Interferometric SAR Statistics

The underlying statistical model for a fully developed speckle field is that of a circular complex Gaussian random process that yields the M-dimensional Gaussian probability density function (PDF) [21]

$$p(\mathbf{g}|\mathbf{C}) = \frac{1}{\pi^M |\mathbf{C}|} \exp(-\mathbf{g}^H \mathbf{C}^{-1} \mathbf{g}) \tag{14}$$

where $\mathbf{C}$ is the covariance matrix. A special case of interest is $M = 2$ for InSAR, which leads to the simplified form for the joint PDF of $g(I_1, I_2, \phi)$

$$p(I_1, I_2, \phi) = \frac{1}{16\pi^2 \sigma^4 (1 - \mu^2)}$$
$$\times \exp \left[-\frac{I_1 + I_2 - 2\sqrt{I_1 I_2}\mu \cos(\phi - \psi)}{2\sigma^2(1 - \mu^2)} \right] \tag{15}$$

where I_1 and I_2 are the intensities of two coregistered SAR images, and it has been assumed that $\langle I_1 \rangle = \langle I_2 \rangle = 2\sigma^2$. ϕ is the noisy interferometric phase. By imposing a scale-invariant similarity criterion, the weight is set as a function of likelihood

$$\mathbf{w}(i_s, j_s) = \prod_m p(I_{1,m}, I_{2,m}, \phi_m)^{1/h} \tag{16}$$

where h is a filtering parameter, the same as in [20]. By applying the maximum likelihood estimation of (15) derived in [22], the estimated parameters can be formulated as

$$\hat{\psi}_{wmle} = -\arg\left(\sum_s \mathbf{w}_s \mathbf{g}_{1,s} \mathbf{g}_{2,s}^*\right) \tag{17}$$

$$\hat{\mu}_{wmle} = \frac{2\sum_s \mathbf{w}_s |\mathbf{g}_{1,s}||\mathbf{g}_{2,s}|}{\sum_s \mathbf{w}_s (|\mathbf{g}_{1,s}|^2 + |\mathbf{g}_{2,s}|^2)} \tag{18}$$

$$\hat{\sigma^2}_{wmle} = \frac{\sum_s \mathbf{w}_s (|\mathbf{g}_{1,s}|^2 + |\mathbf{g}_{2,s}|^2)}{4\sum_s \mathbf{w}_s}. \tag{19}$$

IV. EFFICIENT ALGORITHM FOR SOLVING THE NLCS MODEL

NL filtering and sparse reconstruction algorithms are usually computationally expensive and are difficult to extend to large scales. In this section, we introduce an approach for solving the NLCS model, which can retain the SR power of the standard basis pursuit denoising solver and considerably speed up the processing for matrix $\mathbf{R}$ of the random Fourier transform, as used in SL1MMER.

A. Optimized Parallelization of Nonlocal Process

Note that pixels outside searching windows do not contribute to the value of the central pixel in an NL process. This property allows us to separate the image into independent disjoint pieces and process them in parallel, as it is done in domain decomposition schemes. In [23], we proposed a sophisticated and optimized parallelization scheme for NL processing. A message passing interface was adopted for NL processes, enabling us to use thousands of cores for large-scale processing. The bottleneck of this process is the communication between cores. We introduced a synchronized communication scheme to avoid the bottleneck and the speedup increased dramatically with the increase in the number of cores.

B. Fast and Accurate Solver for L_1 Least Squares Minimization

In SL1MMER, the second-order method, primal–dual interior-point method, with self-dual embedding technique was adopted to solve the second-order cone program, which is extremely expensive with respect to computation. The algorithm proposed in [24], "randomized blockwise proximal gradient," splits the unconstrained optimization problems into two parts, the convex differentiable part and the convex nondifferentiable part, leading to the so-called proximal gradient (PG) method. The iterative approach to solve (12) can be written as

$$\boldsymbol{\gamma}^{k+1} = \arg\min\left(\langle \nabla f(\boldsymbol{\gamma}^k), \boldsymbol{\gamma} - \boldsymbol{\gamma}^k \rangle + \frac{1}{2\alpha_k} \|\boldsymbol{\gamma} - \boldsymbol{\gamma}^k\|_2^2 + r(\boldsymbol{\gamma})\right) \tag{20}$$

where f is $\|\mathbf{R}\boldsymbol{\gamma} - \mathcal{N}(\mathbf{g})\|_2^2$ and ∇f is the partial gradient of function f. The PG formulation is

$$\boldsymbol{\gamma}^{k+1} = \text{prox}_{\alpha_k r}(\boldsymbol{\gamma}^k - \alpha_k \nabla f(\boldsymbol{\gamma}^k)) \tag{21}$$

where $\alpha_k > 0$ is the step size and can be constant or determined by line search. For $r(\boldsymbol{\gamma}) = \|\boldsymbol{\gamma}\|_1$, the proximal operator can be chosen as soft-thresholding.

After applying Nesterov's acceleration scheme and block coordinate techniques, (20) can be written as

$$\boldsymbol{\gamma}_{i_k}^{k+1} = \arg\min\left(\langle \nabla f_{i_k}(\boldsymbol{\gamma}_{i_k}^k), \boldsymbol{\gamma}_{i_k} - \boldsymbol{\gamma}_{i_k}^k \rangle\right.$$
$$\left. + \frac{1}{2\alpha_{i_k}^k} \|\boldsymbol{\gamma}_{i_k} - \boldsymbol{\gamma}_{i_k}^k\|_2^2 + r_{i_k}(\boldsymbol{\gamma})\right) \tag{22}$$

where i_k is the index of a block. The choice of the update index i_k for each iteration is crucial for good performance. Often, it is easy to switch index orders. However, the choice of index affects convergence, possibly resulting in faster convergence or divergence. In this paper, we choose a randomized variants scheme, whose strengths include less memory consumption, good convergence performance, and empirical avoidance of the local optima. The block index i_k is chosen randomly following the probability distribution given by the vector:

$$P_{i_k} = \frac{L_{i_k}}{\sum_{j=1}^J L_j}, \quad i_k = 1, \ldots, J \tag{23}$$

where L_{i_k} is the Lipschitz constant of $\nabla_{i_k} f(\mathbf{x})$, the gradient of $f(\mathbf{x})$ with respect to the i_kth group (in our case $L = \|\mathbf{R}^T \mathbf{R}\|$). However, setting $\alpha_k = 1/L$ usually results in very small step sizes; hence, the time step α_k is adaptively chosen by using the backtracking line search method.
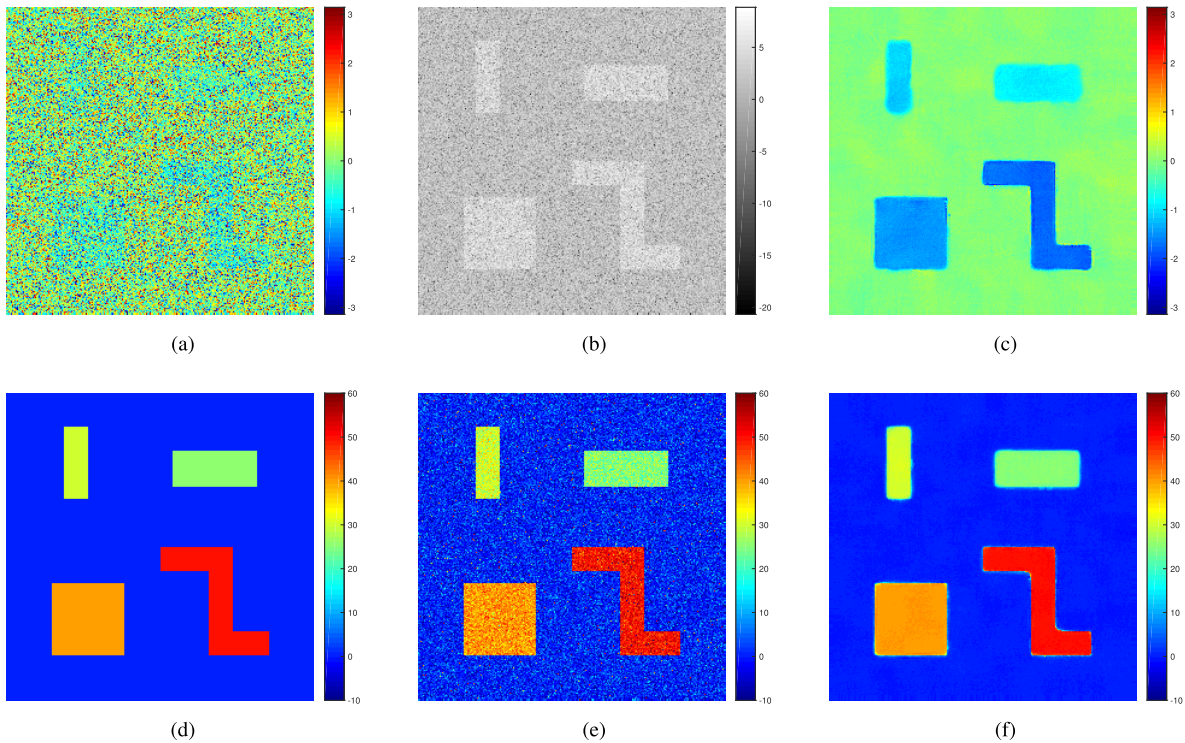
Fig. 3. Simulated data with SNR = 3 dB. (a) One interferogram. (b) Corresponding amplitude. (c) Filtered interferogram. (d) Ground truth of height. (e) Reconstructed height by TomoSAR. (f) Reconstructed height by NLCS-TomoSAR.

TABLE I
PARAMETERS OF NUMERICAL SIMULATION

Shape	Height (meter)	Size (pixels)
Top left	30 m	60×20
Top right	25 m	30×70
Bottom left	40 m	60×60
Bottom right	50 m	$20 \times 50 \times 2 + 20 \times 60$

V. NUMERICAL RESULTS WITH SIMULATED DATA

In this section, we perform our proposed NLCS approach on simulated data. The simulated complex data have been generated from the height profile and different SNR values. An urban-like scene has been generated with rectangular geometric shapes.

The characteristics of the profile and the scene are reported in Table I. The noise-free phase of each interferogram was calculated by using a realistic TerraSAR-X baseline distribution with 29 interferograms

$$\phi_{sim} = \frac{4\pi b_n h_s}{\lambda r \sin \theta_{inc}}. \tag{24}$$

Two stacks of complex data were generated with SNR = 3, −8 dB.

Fig. 3 shows the example of one interferogram and its corresponding amplitude, the interferogram after NL filtering, the ground truth of the buildings' height, the reconstructed height by TomoSAR, and the reconstructed height by NLCS-TomoSAR. As is apparent, the estimation of height by the original TomoSAR can give an acceptable result when the SNR is relatively high. Compared with the original TomoSAR,

NLCS-TomoSAR shows a more accurate result and small loss of resolution at the edges.

Fig. 4 shows the result of the same configuration as in Fig. 3, but with a different SNR (= −8 dB). It can be seen that the interferogram in Fig. 4(a) is strongly blurred and the pattern cannot be easily recognized. After applying the NL filter, the structure of the buildings is visible in Fig. 4(c). The height estimation produced very noisy estimates by TomoSAR for low SNR and the accuracy of the estimates is quite low. In contrast, the estimates of the height by NLCS are extremely good. There is only resolution loss at the edges due to very low SNR.

Fig. 5 presents the slice of height profile along with the black solid line for both SNR = 3 and −8 dB. The black line in Fig. 5(b) and (c) is the ground truth of the height profile and the red dots are the height in each pixel estimated by NLCS-TomoSAR. It can be seen that, for relatively high SNR (3 dB), the spatial bias is not notable and variance is quite small. The standard deviation is 0.23 and 0.24 m, and the mean error is 0.21 and 0.15 m for shapes 1 and 2, respectively. For the low SNR case, both the spatial bias and the variance increase. The detailed results are shown in Table II.

VI. PRACTICAL DEMONSTRATION WITH TERRASAR-X DATA

In this section, we evaluate the performance of the proposed method using the real TerraSAR-X data.

A. Data Description

The test area in this paper is the headquarters of the German Railway, the "Deutsche Bahn" (DB) in Munich. We chose
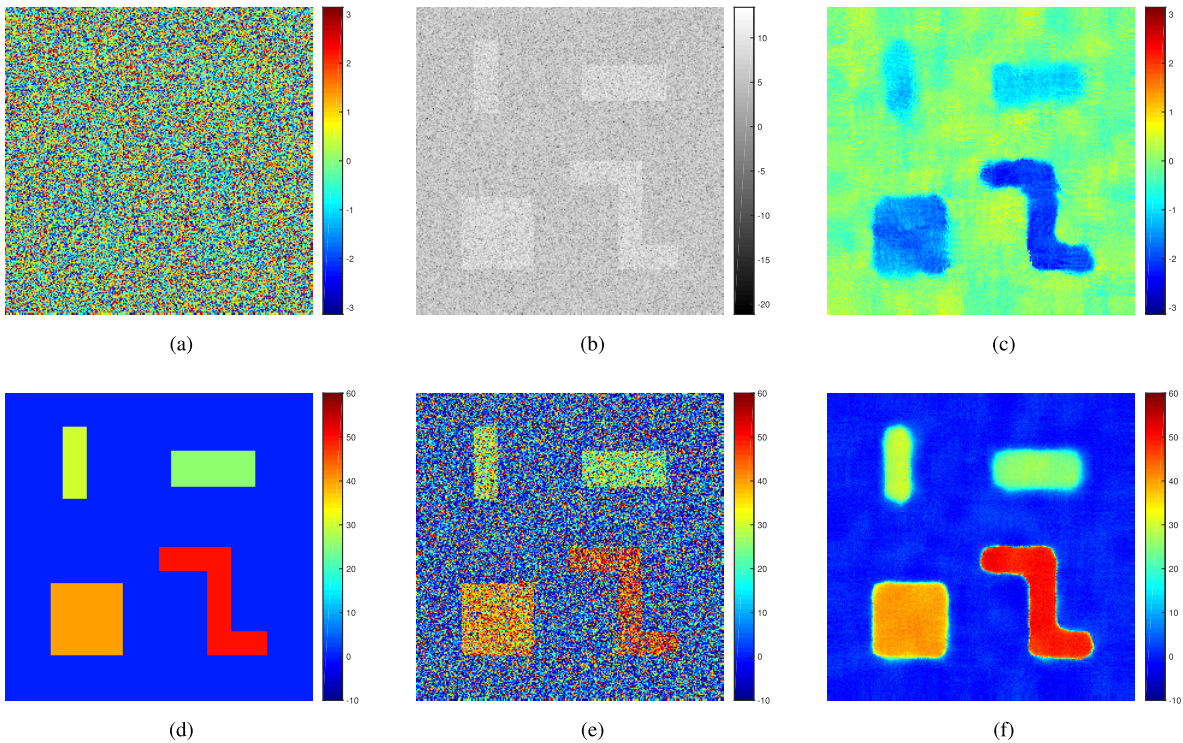
Fig. 4. Simulated data with SNR = −8 dB. (a) One interferogram. (b) Corresponding amplitude. (c) Filtered interferogram. (d) Ground truth of height. (e) Reconstructed height by TomoSAR. (f) Reconstructed height by NLCS-TomoSAR.
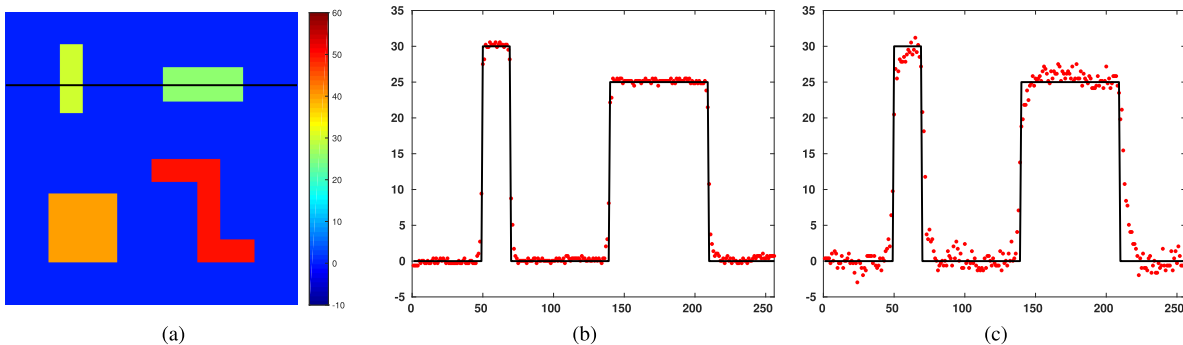

Fig. 5. Height profile of reconstructed result. (a) Position of height profile. (b) Height profile of the case (SNR = 3 dB). (c) Height profile of the case (SNR = −8 dB). Black solid line: ground truth of the height. Red dots: estimated height by the proposed method.

TABLE II
STATISTICS OF HEIGHT ESTIMATION

	Ground Truth	Mean Value	Standard Deviation
Shape 1 (3dB)	30 m	30.21 m	0.23 m
Shape 2 (3dB)	25 m	25.15 m	0.24 m
Shape 1 (-8dB)	30 m	29.21 m	1.44 m
Shape 2 (-8dB)	25 m	25.43 m	1.16 m

TABLE III
TERRASAR-X ACQUISITION PARAMETERS

r	λ	θ_{inc}	Δb
704 km	3.1 cm	39.36°	254.07 m

TerraSAR-X high-resolution spotlight data with a slant-range resolution of 0.6 m and an azimuth resolution of 1.1 m, which consists of 64 interferograms in one stack acquired with a range bandwidth of 300 MHz. The elevation aperture size Δb is about 254.07 m. The detailed parameters of TerraSAR-X acquisition are shown in Table III. The preprocessing, including atmosphere phase screen correction, was performed by the German Aerospace Center PSI-GENESIS system on a persistent scatterer network of high-SNR pixels containing only single scatterers [25]. The SL1MMER algorithm with a Bayesian information criterion [26] as the model selection scheme was applied to each pixel of the test area.

B. Experimental Results

In order to compare the performance of different algorithms, we extracted two new stacks from the original 64 images with 7 images and 14 images, respectively.
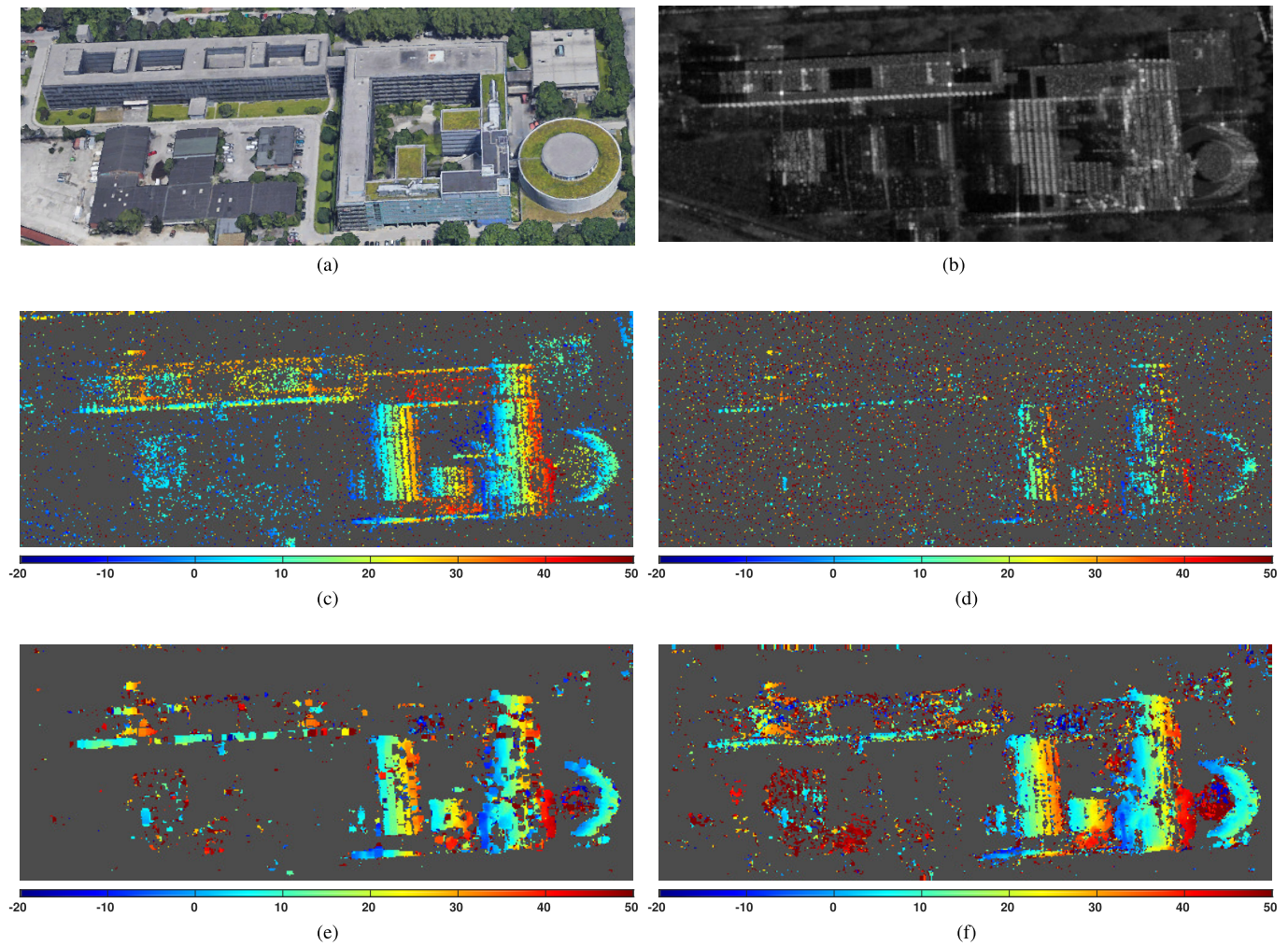
(a)

(b)

(c)

(d)

(e)

(f)

Fig. 6. Test building: DB Headquarters in Munich. (a) Optical image (Copyright Google). (b) Mean amplitude. (c) Elevation estimated by SL1MMER with 64 images. (d) Elevation estimated by SL1MMER with seven images. (e) Elevation estimated by Boxcar + SL1MMER with seven images. (f) Elevation estimated by proposed method with seven images.
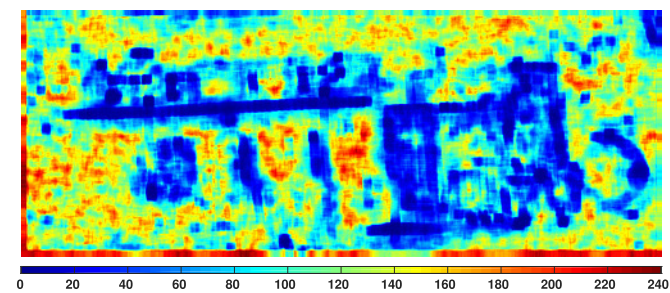

Fig. 7. Averaged equivalent number of looks by the NL filter for the test site.

The corresponding test area of the optical image and the mean amplitude of InSAR stack are shown in Fig. 6(a) and (b), respectively. Fig. 6(c) shows the elevation estimated by the SL1MMER approach with 64 images and Fig. 6(d) shows the elevation estimated by the SL1MMER approach with seven images. Note that the reconstructed result is not satisfactory with only seven images. The estimated elevation exhibits strong noise due to the small number of images.

In addition, the successfully reconstructed elevation is significantly less than the reconstructed result with 64 images. As a comparison, we show the reconstructed result with a boxcar filter. Fig. 6(e) is the elevation estimated by a boxcar filter with a window size of 5. It is clear that the loss of resolution is dramatic compared with the original TomoSAR. Boxcar filtering blurs edges and small structures present in the images. As can be seen in Fig. 6(e), the proposed method can obtain an extraordinarily good estimated result with only seven images. In contrast with the boxcar filter, the building structures are retrieved by our method, both in terms of shapes and elevations, without notable resolution loss.

Fig. 7 shows the averaged equivalent number of looks by an NL filter for the test site. From this figure, we can see that the elevation is estimated in a spatially adaptive manner. The number of looks at the buildings is quite lower than in the homogeneous area, which indicates that pixels chosen by an NL filter should have similar properties, such as similar elevation, reflectivity, scattering characteristics, and so on.

In [15], it was shown that a 90% detection rate of two scatterers with a distance of ρ_s can be achieved, while
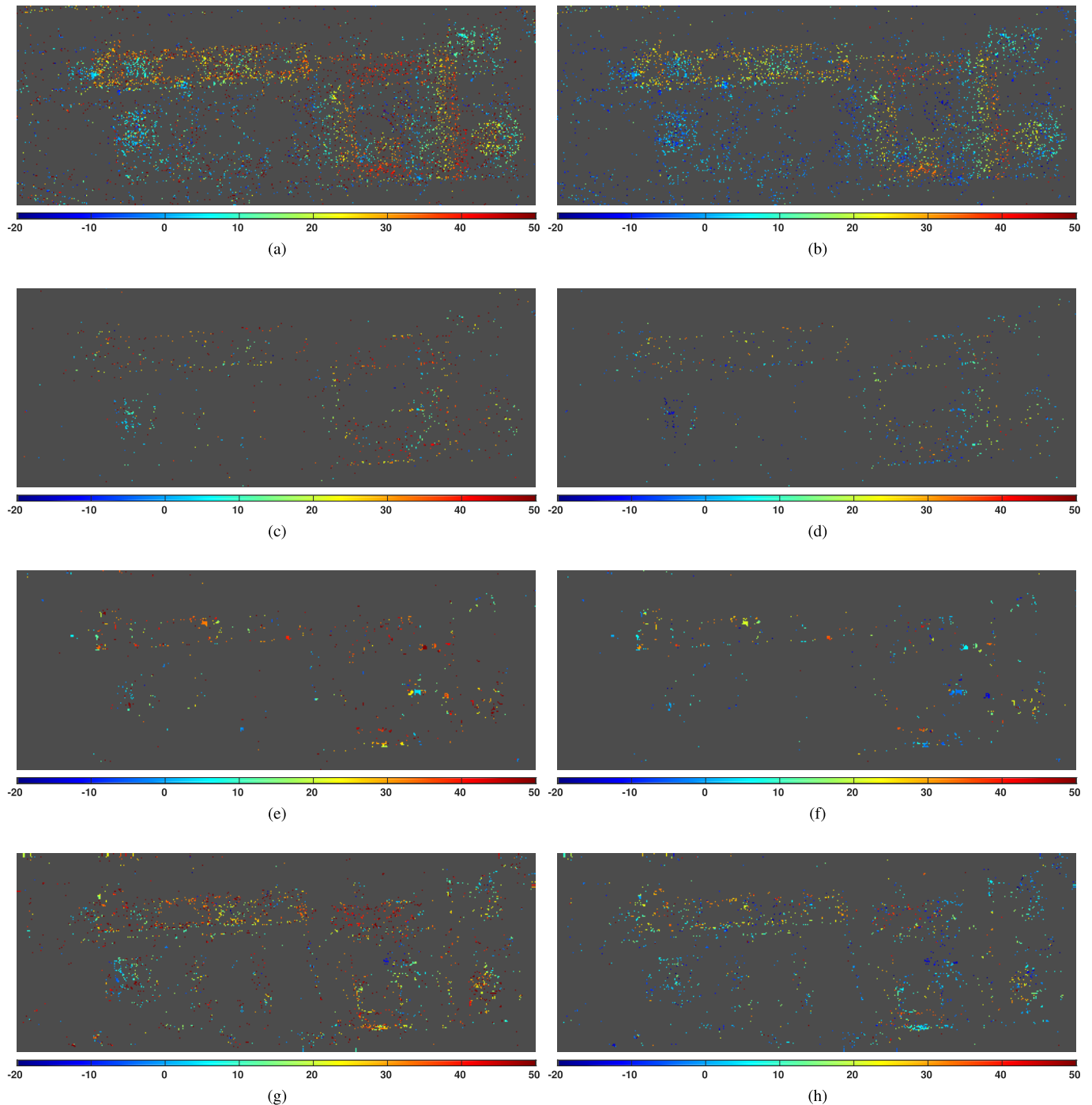
Fig. 8. Test building: DB Headquarter in Munich. Elevation estimates of the separated double scatterers. (a) Top layer (64 images + TomoSAR), mostly caused by returns from building facade. (b) Ground layer (64 images + TomoSAR), mostly caused by returns from ground structures. (c) Top layer (14 images + TomoSAR). (d) Ground layer (14 images + TomoSAR). (e) Top layer (14 images + Boxcar-TomoSAR). (f) Ground layer (14 images + Boxcar-TomoSAR). (g) Top layer (14 images + NLCS-TomoSAR). (h) Ground layer (14 images + NLCS-TomoSAR).

17 acquisitions are needed when the amplitude of reflectivity of one scatterer is twice that of another scatterer. In the real scenario, this is most often the case when one scatterer sits on the facade or roof and another is on the ground. Fig. 8 shows the elevation estimates of separated double scatterers. Fig. 8(a), (c), (e), and (g) shows the top layer of scatterers reconstructed by TomoSAR with 64 images, TomoSAR with 14 images, Boxcar-TomoSAR with

14 images, and NLCS-TomoSAR with 14 images, respectively. Fig. 8(b), (d), (f), and (h) shows the ground layer of scatterers for the four cases. It is clear that TomoSAR with 64 images in Fig. 8(a) and (b) is proficient at reconstructing double scatterers, i.e., a top layer mostly caused by reflections from the facade of the building and a ground layer caused by reflections from lower buildings and ground infrastructures. As can be seen, keeping a similar SNR, TomoSAR with
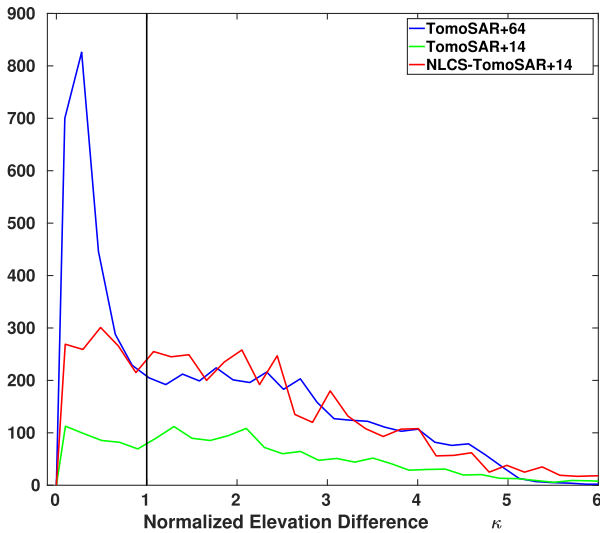
Fig. 9. Histogram of double scatterers' elevation differences using TomoSAR with 64 images (blue line), TomoSAR with 14 images (green line), and NLCS-TomoSAR with 14 images (red line).

14 images produces a very low detection rate for double scatterers in Fig. 8(c) and (d). After applying the boxcar filter, the detection rate increases a little bit, but the resolution decreases a lot. In contrast, the number of double scatterers detected by the NLCS approach with small stacks is comparable to TomoSAR with large stacks, and its resolution loss is not obvious.

The histogram of the double scatterers' elevation differences using TomoSAR with 64 images (blue line), TomoSAR with 14 images (green line), and our NLCS-TomoSAR with 14 images (red line) is shown in Fig. 9. The normalized distance is defined as

$$\kappa = \frac{s}{\rho_s}. \tag{25}$$

Note that the CS-based SL1MMER reconstruction with the large stack has impressive SR capability, i.e., many of the double scatterers with $\kappa < 1$ are detected. However, when the number of interferograms decreases, the performance of double-scatterer detection decreases accordingly. In contrast, the proposed NLCS-TomoSAR method with only 14 images obtains the same result as the large stack reconstruction, at least in the non-SR regime ($\kappa > 1$). In the SR region, it is still much better than the standard method but falls short of the 64-stack reconstruction. We assume that this is caused by the averaging effect of NL filtering. Since target responses of different amplitudes and different sub-pixel positions are averaged, the resulting amplitudes may be slightly compromised. The baseline-dependence of amplitude, however, is an important indicator of double scatterers.

VII. Conclusion

In this paper, we propose a novel framework for TomoSAR with a minimum number of acquisitions in order to obtain a fast and accurate estimation of elevation without any *a priori* knowledge. We evaluated the performance of the proposed

NLCS algorithm with simulated and real data. Experiments using the simulated data illustrate that the proposed method can give excellent height estimation for different SNRs without notable resolution distortion in comparison to the state-of-the-art methods such as SL1MMER. Moreover, using only seven SAR images over the test site Munich, it is practically demonstrated that NLCS can achieve an accurate height estimation while preserving detailed structures. Furthermore, due to the increased SNR, a remarkable layover separation capability of NLCS can be observed.

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, Sep. 2000.

[2] 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, Oct. 2002.

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

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

[5] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[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. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 10, pp. 3743–3756, Oct. 2018.

[8] X. X. Zhu, Y. Wang, S. Montazeri, and N. Ge, "A review of ten-year advances of multi-baseline SAR interferometry using TerraSAR-X data," *Remote Sens.*, vol. 10, no. 9, p. 1374, Aug. 2018.

[9] X. Zhu and R. Bamler, "Very high resolution SAR tomography via compressive sensing," in *Proc. FRINGE Workshop*, Frascati, Italy, 2009, pp. 1–7.

[10] 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.

[11] 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.

[12] E. Aguilera, M. Nannini, and A. Reigber, "Wavelet-based compressed sensing for SAR tomography of forested areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 12, pp. 5283–5295, Dec. 2013.

[13] P. Ma, H. Lin, H. Lan, and F. Chen, "On the performance of reweighted L_1 minimization for tomographic SAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 895–899, Apr. 2015.

[14] 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.

[15] 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.

[16] X. X. Zhu, N. Ge, and M. Shahzad, "Joint sparsity in SAR tomography for urban mapping," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.

[17] 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.

[18] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE CVPR*, San Diego, CA, USA, Jun. 2005, pp. 60–65.

[19] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[20] C.-A. Deledalle, L. Denis, and F. Tupin, "NL-InSAR: Nonlocal interferogram estimation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 4, pp. 1441–1452, Apr. 2011.

[21] J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications*. Englewood, CO, USA: Roberts Company Publishers, 2007.

[22] M. S. Seymour and I. G. Cumming, "Maximum likelihood estimation for SAR interferometry," in *Proc. IEEE IGARSS*, Aug. 1994, pp. 2272–2275.

[23] Y. Shi, X. X. Zhu, and R. Bamler, "Optimized parallelization of nonlocal means filter for image noise reduction of InSAR image," in *Proc. IEEE ICIA*, Aug. 2015, pp. 1515–1518.

[24] 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.

[25] N. Adam, M. Eineder, N. Yague-Martinez, and R. Bamler, "High resolution interferometric stacking with TerraSAR-X," in *Proc. IEEE IGARSS*, vol. 2, Jul. 2008, pp. II-117–II-120.

[26] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, no. 2, pp. 461–464, Mar. 1978.

Yilei Shi (M'18) received the Diploma degree in mechanical engineering from the Technische Universität München (TUM), Munich, Germany, in 2010.

He is currently a Research Associate with the Chair of Remote Sensing Technology, TUM. His research interests include fast solver and parallel computing for large-scale problems, advanced methods on SAR and InSAR processing, machine learning and deep learning for variety of data sources, such as SAR, optical images, and medical images, and PDE-related numerical modeling and computing.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the M.Sc., Dr.-Ing., and Habilitation degrees in signal processing from the Technische Universität München (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She is currently a Professor for Signal Processing in Earth Observation, TUM, and the German Aerospace Center (DLR), Weßling, Germany; the Head of the Department of EO Data Science, DLR; and the Head of the Helmholtz Young Investigator Group, SiPEO at DLR and TUM. She was a Guest Scientist or a Visiting Professor at the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, the University of Tokyo, Tokyo, Japan, and the University of California at Los Angeles, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. Her research interests include remote sensing and earth observation, signal processing, machine learning, and data science, with a special application focus on global urban mapping.

Dr. Zhu is a member of the Young Academy (Junge Akademie/Junges Kolleg) at the Berlin-Brandenburg Academy of Sciences and Humanities, the German National Academy of Sciences Leopoldina, and the Bavarian Academy of Sciences and Humanities. She is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Richard Bamler (M'95–SM'00–F'05) received the Diploma degree in electrical engineering, the Ph.D. degree in engineering, and the Habilitation degree in signal and systems theory from the Technische Universität München (TUM), Munich, Germany, in 1980, 1986, and 1988, respectively.

He was with TUM from 1981 to 1989, where he was involved in optical signal processing, holography, wave propagation, and tomography. He joined the German Aerospace Center (DLR), Weßling, Germany, in 1989, where he is currently the Director of the Remote Sensing Technology Institute.

In 1994, he was a Visiting Scientist at the Jet Propulsion Laboratory, Pasadena, CA, USA, where he is involved in the preparation of the SIC-C/X-SAR missions, and in 1996, he was a Guest Professor at the University of Innsbruck, Innsbruck, Austria. Since 2003, he has been holding a full professorship in remote sensing technology at TUM as a double appointment with his DLR position. His teaching activities include university lectures and courses on signal processing, estimation theory, and SAR. Since he joined DLR, he, his team, and his institute have been working on SAR and optical remote sensing, image analysis and understanding, stereo reconstruction, computer vision, ocean color, passive and active atmospheric sounding, and laboratory spectrometry. They were and are responsible for the development of the operational processors for SIR-C/X-SAR, SRTM, TerraSAR-X, TanDEM-X, Tandem-L, ERS-2/GOME, ENVISAT/SCIAMACHY, MetOp/GOME-2, Sentinel-5P, Sentinel-4, DESIS, and EnMAP. His research interests include algorithms for optimum information extraction from remote sensing data with an emphasis on SAR. This involves new estimation algorithms, such as sparse reconstruction, compressive sensing, and deep learning.