

Building 3-D Reconstruction With a Small Data Stack Using SAR Tomography

Hongliang Lu, *Student Member, IEEE*, Heng Zhang , *Member, IEEE*, Yunkai Deng, *Member, IEEE*, Jili Wang , and Weidong Yu, *Member, IEEE*

Abstract—Synthetic aperture radar (SAR) tomography (TomoSAR) has been well established for the 3-D reconstruction of urban buildings. Many methods have been proposed in the literature for TomoSAR inversion. These methods usually require fairly large data stacks (more than 20 images) for reliable reconstruction. Hence, they cannot be applied to the 3-D reconstruction of urban buildings using Gaofen-3 (GF-3) data directly, because there are few images available on average in each city. This article proposes a novel workflow for the 3-D reconstruction of high-rise buildings using small data stacks. In this workflow, we combine the methods of contour line extraction (CLE) and reference-elevation multilooking relaxation (RM-RELAX) for the 3-D reconstruction of high-rise buildings. The CLE method extracts contour lines of buildings relying on a data-driven approach and does not need to rely on any external data, which can provide prior knowledge for subsequent processing. The RM-RELAX method is an extension of RELAX, which can obtain more precise 3-D inversion with multilooking and reject outliers with the reference elevation (which can be obtained by the prior knowledge of contour lines) of each pixel. The applicability of this workflow is demonstrated by using simulated data and real data with six SAR images acquired by the GF-3 satellite over an area in Beijing, China.

Index Terms—Contour line extraction (CLE), Gaofen-3 (GF-3), reference-elevation multilooking relaxation (RM-RELAX), synthetic aperture radar (SAR) tomography (TomoSAR).

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) tomography (TomoSAR) has been widely used in the 3-D reconstruction of urban buildings [1]–[5]. It uses the coherent information of the SAR images of repeat-pass acquisitions to reconstruct reflectivity profiles of point-like and volumetric scatterers along the elevation direction. There are many methods that have been proposed for TomoSAR imaging, such as, beam forming [6], truncated singular value decomposition [3], Capon [7],

multiple signal classification (MUSIC) [8], relaxation algorithm (RELAX) [9], and compressive sensing (CS) [5]. Limited by the span of baselines, the elevation resolution of TomoSAR is generally on the order of tens of meters, which is much worse than the azimuth and range resolution. Therefore, TomoSAR commonly requires superresolution processing methods. Capon, MUSIC, RELAX, and CS allow superresolution imaging, and the CS method has the best superresolution capability [10].

The main disadvantage of these methods is that they require large data stacks (more than 20 images) for reliable reconstruction. They are usually not directly applicable to TomoSAR inversion using small data stacks. A small data stack has the features of a short time of data acquisition and fast updates to urban building elevations. However, it has only a few SAR images and even large baseline intervals. These characteristics cause a severe ambiguity issue and many strong aliasing targets and even false targets in elevation imaging, which will make 3-D reconstruction fail. In [10], the neighborhood-based methods, distributed CS (DCS) and multilooking method to CS (MCS) were demonstrated to perform better than the standard CS when dealing with a small data stack using simulated data. Zhu *et al.* [11] proposed a joint sparsity concept to obtain precise TomoSAR inversion with only six images by combining *a priori* knowledge of high-rise buildings to the estimation. However, this method can only be used to reconstruct buildings at places where geographic information system (GIS) data are available, because precise geometric prior knowledge is an essential factor for its application. In [12] and [13], the authors integrated nonlocal (NL) estimation into TomoSAR inversion to increase the signal-to-noise ratio (SNR). The results show that a reasonable reconstruction of buildings with only seven images is feasible. However, this method will cause clusters of outliers. All of these methods use the idea of a multilooking method. The multilooking method is a simple and effective method to suppress noise, and its benefits have been demonstrated in many studies [14]–[17]. Therefore, we will also take the benefits of the multilooking method to process the 3-D reconstruction of buildings with a small data stack.

In this article, a novel workflow is proposed to handle small data stacks. The main contributions are as follows.

- 1) We propose a contour line extraction (CLE) method to extract the contour lines of high-rise buildings as prior knowledge of subsequent processing.
- 2) We extend RELAX to the reference-elevation multilooking relaxation (RM-RELAX) method by combining the

Manuscript received December 31, 2019; revised March 5, 2020, April 14, 2020, and April 26, 2020; accepted May 13, 2020. Date of publication May 19, 2020; date of current version June 4, 2020. This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFB0502700 and in part by the National Natural Science Foundation under Grant 61901443. (*Corresponding author: Heng Zhang.*)

Hongliang Lu is with the Space Microwave Remote Sensing System Department, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China, and also with the School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100039, China (e-mail: luhongliang16@mails.ucas.edu.cn).

Heng Zhang, Yunkai Deng, Jili Wang, and Weidong Yu are with the Space Microwave Remote Sensing System Department, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China (e-mail: caszhmail@163.com; ykdeng@mail.ie.ac.cn; wangjili13@mails.ucas.ac.cn; ywd@mail.ie.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2020.2995503

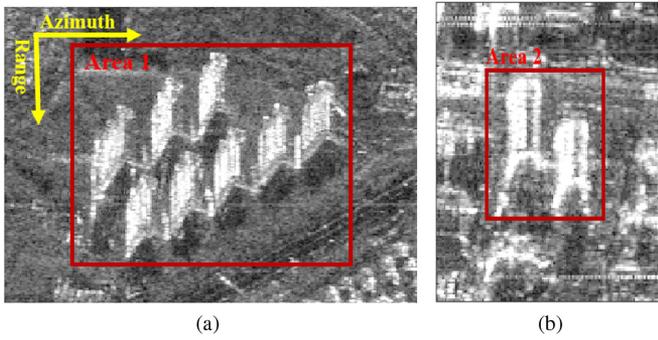


Fig. 1. SAR image of interesting areas acquired by the GF-3 satellite.

multilooking method and reference elevation (which can be obtained by the prior knowledge of contour lines) of each pixel, which can obtain more precise estimates and reduce the occurrence of outliers in the 3-D reconstruction of high-rise buildings.

- 3) The performance of RM-RELAX is systematically evaluated in terms of elevation estimation accuracy and the detection rate of separating overlaid double scatterers with simulated data. Compared to the performance of the RELAX and MCS methods, that of the proposed method is better and evident for all of the aforementioned quantitative metrics.
- 4) The first tomographic reconstruction using China's spaceborne GF-3 data is presented. The applicability of the proposed workflow is demonstrated in practice.

The rest of this article is organized as follows. In Section II, the characteristics of GF-3 data are introduced. Section III mainly discusses the CLE of high-rise buildings by using the CLE method. In Section IV, the principle and implementation of the proposed RM-RELAX method are explained. In Section V, experiments using simulated data and real data are presented. Finally, Section VI concludes this article.

II. DATA STACK

The real SAR data stack used in this article is acquired by China's spaceborne GF-3 satellite. It has six fine stripmap I mode (FSM_I) GF-3 images acquired in ascending mode, from March 2017 to May 2018 over the metropolitan area of Beijing, China. The area of interest is shown in Fig. 1. The spatial and temporal baseline configurations of these six images are shown in Fig. 2(a), and the basic parameters of GF-3 are shown in Table I. The span of the spatial baselines is 2311.5 m, and the maximum baseline interval is 921.29 m. The span of the temporal baseline is 406 days. Fig. 2(b) shows the average coherence of Area 1 in Fig. 1(a). This demonstrates that the decorrelation of these data stacks is important due to the large interval of spatial and temporal baselines.

III. CLE METHOD

The contour lines of high-rise buildings are important prior knowledge in this article. Contour lines can provide the pixel

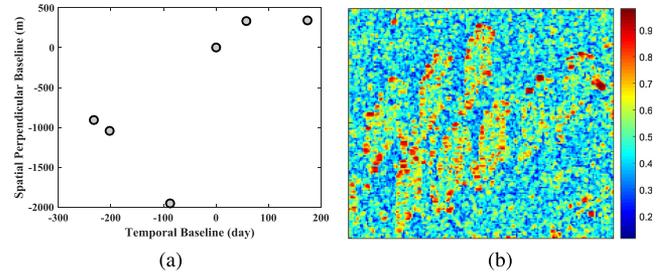


Fig. 2. (a) Spatial and temporal baseline configurations and the (b) average coherence of GF-3 data stacks.

TABLE I
PARAMETERS OF GF-3

Parameters	Quantity
Slant-range (r_0)	900 km
Wavelength (λ)	5.55 cm
Incident angle (θ)	35°
Rayleigh resolution of elevation (ρ_s)	10.75 m
Maximum interval of baseline (Δb_{\max})	921.29 m

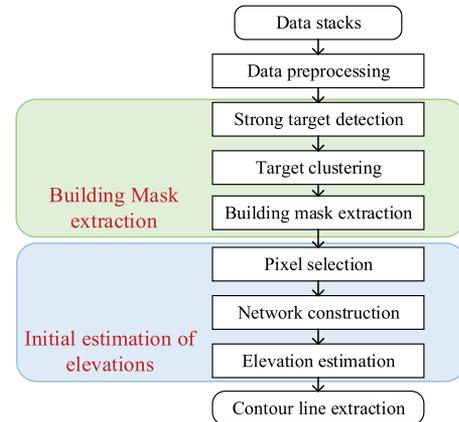


Fig. 3. Processing sequence of the CLE method.

index for pixel grouping in subsequent processing. In [11], a method using GIS data was proposed. This method can obtain accurate contour lines through mapping relationships. It is not suitable for buildings without GIS data. This section describes a data-driven CLE method, which can extract contour lines without relying on any external data. The processing sequence of this method is shown in Fig. 3.

To obtain a more intuitive understanding of the CLE method, we provide the processing results using GF-3 SAR data stacks. We select the high-rise buildings in Area 1 marked by the red box in Fig. 1(a) for this test.

A. Data Preprocessing

The geometry of TomoSAR imaging is shown in Fig. 4. The data preprocessing includes coregistration of images, amplitude calibration, and deramping. After these operations, the signal

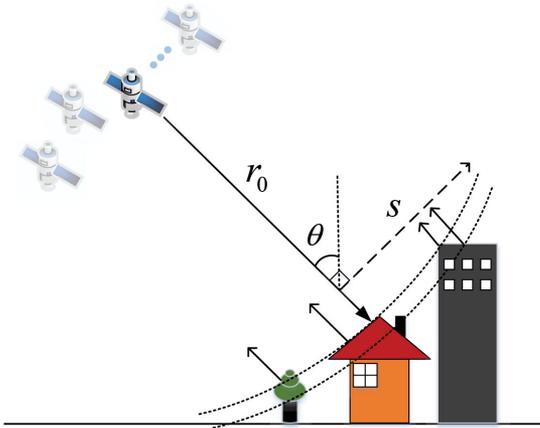


Fig. 4. Geometry of TomoSAR imaging.

model [3] can be written as

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

where $g_n (n = 1, 2, \dots, N)$ is the complex-valued measurement at an azimuth-range pixel of the image, $\gamma(s)$ is the reflectivity function along elevation s with an extension of ΔS , $\xi_n = 2b_n/\lambda r_0$ is the spatial frequency depending on the orthogonal baseline b_n , λ is the wavelength of the radar signal, and r_0 denotes the range between the radar and the observed target.

Considering the existence of noise e , the discrete TomoSAR system can be rewritten as

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \mathbf{e} \quad (2)$$

where \mathbf{g} is the measurement vector with N single-look complex observations, $\boldsymbol{\gamma}$ is the discrete and sparse reflectivity vector along the elevation direction, and \mathbf{R} is the $N \times L$ (L is the number of discrete samplings in the elevation direction) steering matrix with $R_{nl} = \exp(j2\pi\xi_n s_l)$.

B. Building Mask Extraction

The interesting area where high-rise buildings are located is selected for CLE. The first thing to do is building mask extraction.

First, the constant false alarm rate (CFAR) [18] detection method is used to extract strong (buildings) targets in SAR images. The CFAR is a classic method for bright target detection with some characteristics of simple computation, constant false alarm probability, adaptive threshold, and fast detection of targets from complex backgrounds [19]. The three critical elements of the CFAR are sliding window size, false alarm probability, and detection threshold. The sliding window is shown in Fig. 5. It is a square hollow-stencil sliding window, which is suitable for high-resolution SAR target detection [20]. In this test, considering the size of the high-rise buildings, we set the width of the guard area to 20 pixels. The false alarm probability is set to 10^{-3} . The adaptive detection threshold depends on the false alarm probability and the clutter statistics of the clutter region, and the explicit derivation of the formula can be found in [19].

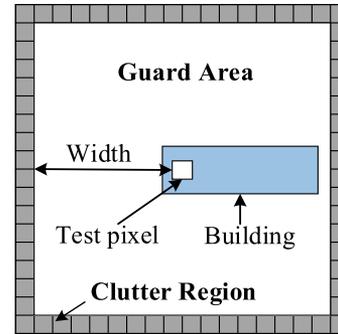


Fig. 5. Sliding window of the CFAR.

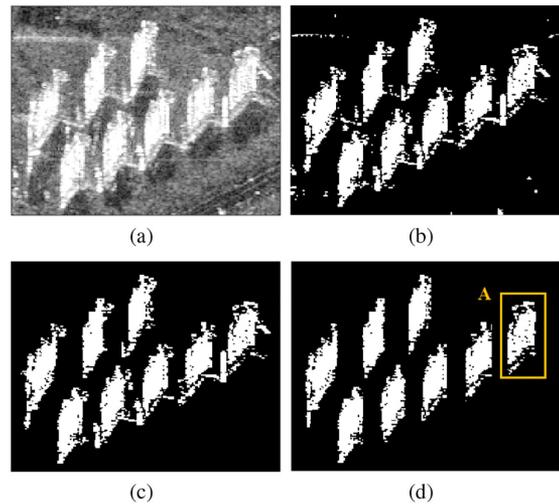


Fig. 6. Building mask extraction. (a) SAR intensity image. (b) Strong target detection. (c) Target clustering. (d) Building mask extraction.

Fig. 6(b) shows the result of detecting strong targets using the CFAR method, and the strong targets are well-detected.

Second, strong targets detected by the CFAR are clustered through the connection relationships among targets. We set up a window with a size of 3×7 pixels in the range-azimuth direction to determine the connection relationship. The target in the center of the window and targets in the window are considered connected and can be clustered into one region. After clustering, all targets are divided into different regions. Smaller regions still exist. These undesired regions will be eliminated. Fig. 6(c) shows the clustering results.

Finally, a neighborhood window is used to extract high-rise buildings. Considering both the spatial resolution and the general size of high-rise buildings, the size of the window is set to 3×31 pixels. If the number of strong pixels in the neighborhood window is more than half of the total number of pixels in the window, this pixel is retained as the pixel of high-rise buildings. Fig. 6(d) shows the result of the building mask extraction.

After the building mask extraction, the coordinate range of the pixels can be obtained in the range-azimuth direction of each building in the SAR images. Then, the contour lines of each building can be extracted separately.

C. Initial Estimation of Elevations

To ensure the accuracy of the elevation estimation, we should first remove the atmospheric phase screen (APS). Since the data stack has only six SAR images, the persistent scatterer interferometry (PS-InSAR) method used to estimate APS will be inaccurate. In [4], a method of constructing a network for TomoSAR with no need for preliminary removal of the APS was introduced. In this step, a modified algorithm according to this method is applied.

First, some reliable pixels (RPs) of the building should be selected to construct a Delaunay triangulation network (DTN) [21], [22]. Amplitude dispersion is a classical criterion for RP selection. However, it is not applicable for our data stack with only six SAR images, so we need to find another criterion to assist in the selection of RPs. As comprehensively investigated in [23], the product of the number of acquisitions and the SNR determines the reconstruction quality, which means that increasing the SNR can reduce the number of acquisitions in TomoSAR inversion. This conclusion can also be confirmed from the Cramer–Rao lower bound (CRLB) of the elevation estimates using tomographic imaging

$$\sigma_s = \frac{\lambda r_0}{4\pi \cdot \sigma_b \cdot \sqrt{2} \cdot N \cdot \text{SNR}} \quad (3)$$

where σ_b is the standard deviation of the baseline. Therefore, we can select pixels with high SNRs (relatively large mean amplitude of six SAR images) as RPs to construct the DTN.

Second, we filter the arcs in the DTN and estimate the relative elevations of RPs. For a pair of connected RPs, we choose one as the start RP and the other as the end RP. As the two RPs are adjacent (in one region of a building), they can be assumed to have the same APS. We can remove the APS for the end RP by subtracting the phase of the start RP. After calibrating the phase, we can use beamforming to estimate the parameters of the end RP. The inversion can be written as [4]

$$\hat{\gamma} = \frac{|\mathbf{r}(s)^H \mathbf{g}_{1,2}|}{\|\mathbf{r}(s)\|_2 \|\mathbf{g}_{1,2}\|_2} \quad (4)$$

where

$$\mathbf{r}(s) = \begin{bmatrix} \exp(j2\pi\xi_1 s) \\ \vdots \\ \exp(j2\pi\xi_N s) \end{bmatrix} \quad (5)$$

$$\mathbf{R} = [\mathbf{r}(s_1), \dots, \mathbf{r}(s_L)] \quad (6)$$

where $\mathbf{g}_{1,2}$ denotes the measurement vector of the end RP after calibrating the phase, $(\cdot)^H$ represents the transpose and conjugate operation, and $\|\cdot\|_2$ is the 2-norm. The tomographic magnitude in (4) has been normalized. The maximum normalized tomographic magnitude $\hat{\gamma}_{rp}$ and relative elevation s_{rp} of the end RP can be extracted by detecting the maximum peak. In the case of PSs, $\hat{\gamma}_{rp}$ is almost equivalent to the ensemble coherence because single PSs normally have low amplitude dispersion [4]. Therefore, a threshold T_M can be set to select PSs. If $\hat{\gamma}_{rp} > T_M$, we consider the two RPs at the arc to be true PSs. Since there are only six images in the GF-3 data stack and the interval of

baselines is large, there is a severe ambiguity issue in elevation imaging, so another threshold T_S should be added to ensure the quality of the relative elevation estimation. If $s_{rp} < T_S$, the estimation of relative elevation is considered reliable. If the estimated parameters of the end RP satisfy both constraints, we preserve the arc. Otherwise, the arc is rejected from the network. A weighted least square (WLS) estimator mentioned in [4] is used to improve the precision of the elevation estimation. The preliminary estimates can be used to perform temporal phase unwrapping. We can reformulate the inversion problem using the unwrapping phase $\Delta\phi$ as follows:

$$\Delta\phi = \mathbf{D}s \quad (7)$$

where

$$\Delta\phi = [\Delta\varphi_1, \dots, \Delta\varphi_N]^T \quad (8)$$

$$\mathbf{D} = [2\pi\xi_1, \dots, 2\pi\xi_N]^T \quad (9)$$

where $(\cdot)^T$ represents the transpose operation. Therefore, it is possible to calculate the final relative elevation s in a continuous searching space with a WLS estimator, i.e.,

$$s = (\mathbf{D}^T \mathbf{W} \mathbf{D})^{-1} \mathbf{D}^T \mathbf{W} \Delta\phi \quad (10)$$

where \mathbf{W} is an $N \times N$ diagonal matrix, and the detailed calculation method is described in [4].

Finally, we integrate the relative elevations through network adjustment and retrieve absolute elevations using one reference pixel by using WLS estimators. There will be some isolated small networks after rejecting the arcs. Before network adjustment, it is necessary to identify the largest connected network because the disconnected network cannot be adjusted together. It is assumed that there are P arcs and Q RPs in the largest connected network, and the integration problem can be formulated as

$$\mathbf{S} = \mathbf{A}\mathbf{X} \quad (11)$$

where \mathbf{S} contains the estimated relative elevations at the arcs; \mathbf{A} is the adjustment matrix consisting of -1 , 0 , and 1 ; and \mathbf{X} contains the absolute elevations of the preserved RPs, i.e.,

$$\mathbf{S} = [s_1, \dots, s_P]^T \quad (12)$$

$$\mathbf{A} = \begin{bmatrix} 1 & \dots & -1 & \dots & 0 & \dots \\ \vdots & & & & & \\ \dots & 1 & 0 & \dots & -1 & \dots \\ \vdots & & & & & \end{bmatrix} \quad (13)$$

$$\mathbf{X} = [x_1, \dots, x_Q]^T \quad (14)$$

where -1 at the row of \mathbf{A} represents the start RP of one arc, and 1 represents the end RP of the arc. As the rank of \mathbf{A} is always $Q - 1$, \mathbf{A} is singular. We can choose one RP in \mathbf{X} to act as the reference pixel for all of the RPs and subtract the corresponding column of the reference pixel in \mathbf{A} so that \mathbf{A} becomes a full-rank matrix. Then, the WLS estimator can be used to integrate the relative elevations [24], i.e.,

$$\mathbf{X} = (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \mathbf{S} \quad (15)$$

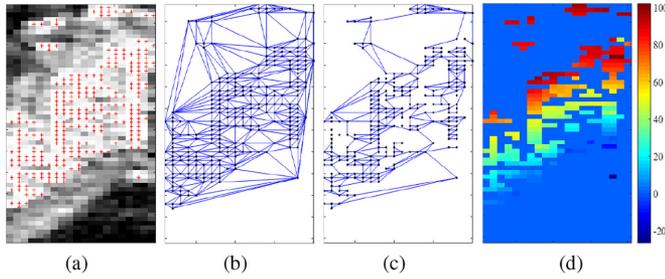


Fig. 7. Initial estimation of the elevation of the building A. (a) Selected RPs, (b) DTN, (c) preserved arcs, and (d) estimated elevations.

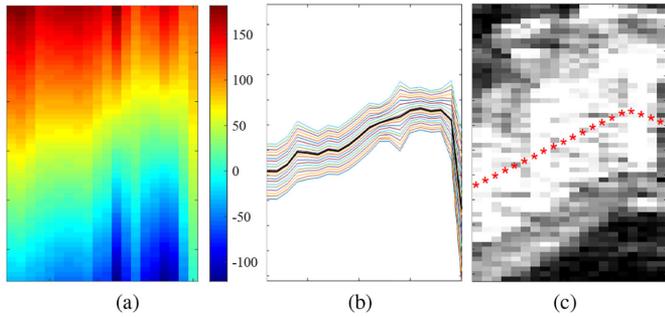


Fig. 8. CLE of building A. (a) Elevations of all pixels. (b) Colored contour lines. (c) Final contour line.

where \mathbf{W} is a $P \times P$ diagonal matrix that normally contains the ensemble coherence (here, the maximum normalized tomographic magnitude $\hat{\gamma}_{rp}$) at each arc as a weight factor to represent the quality of each estimate.

We take building A marked by a yellow box in Fig. 6(d) as an example to show the processing results of the initial estimation of elevations. In this test, we select pixels with mean amplitudes greater than 200 as RPs to construct a DTN, and the results are shown in Fig. 7(a) and (b). We set $T_M = 0.7$ and $T_S = 30$ (the value of T_S depends on the ambiguity issue of the data stack, which will be introduced in Section IV) to reject the unreliable arcs, and then, use network adjustment and the WLS estimator to obtain the initial estimation of the elevations of building A. Fig. 7(c) shows the preserved arcs and Fig. 7(d) shows the estimated elevations of building A.

D. Contour Line Extraction (CLE)

We use the estimated elevations of RPs to extract the contour lines of the building as follows.

1) We calculate the amplitude dispersion value d_{ij} and the mean amplitude value a_{ij} of all RPs using six SAR images, where i and j represent the row (azimuth) and column (range) coordinates, respectively, of RPs within the building mask. We define a_{mj} as the median of all a_{ij} in column j . For each column j , we use the elevations of RPs with $d_{ij} < 0.25$ or $a_{ij} > a_{mj}$ for first-order polynomial fitting to obtain the elevations of all pixels in the column j . Therefore, we can obtain the elevations of all pixels in the building mask, as shown in Fig. 8(a).

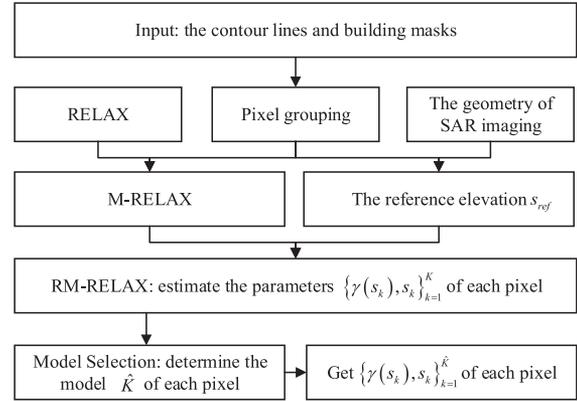


Fig. 9. Procedure of the RM-RELAX method.

2) One pixel with the elevation closest to s_c is selected in each column, and then, these selected pixels will be connected to obtain one contour line. In this test, we choose 20 different elevation values s_c between 25 and 65 m and obtain 20 contour lines, as shown by the colored polylines in Fig. 8(b). Then, we calculate the average row coordinates of the 20 pixels selected in each column to obtain a reliable contour line, as shown by the thicker black line in Fig. 8(b).

3) Considering the regularity of high-rise buildings, such as cubes and cylinders, we use piecewise polynomial fitting on the thicker black contour line to obtain the final contour line of building A, which is shown in Fig. 8(c).

IV. RM-RELAX METHOD

The reference-elevation multilooping relaxation (RM-RELAX) method proposed in this section is an extension of RELAX. It can obtain more precise 3-D inversion with multilooping and reject outliers with reference elevation of each pixel. Fig. 9 shows the procedure of the RM-RELAX method.

A. Standard RELAX-Based TomoSAR

The relaxation (RELAX) algorithm is a high-resolution spectral estimation method based on the nonlinear least-squares (NLS) criterion [25], [26]. The RELAX algorithm does not make any restrictive assumptions on noise. It obtains the estimated values of the parameters by minimizing the nonlinear variance. Even in the background of colored noise, it can obtain the asymptotically effective estimation of each parameter with good robustness.

The standard RELAX-based TomoSAR was described in [27]. Under the principle of least squares, the sparse signal can be estimated by the following equation:

$$\begin{aligned} \{\gamma(s_k), s_k\}_{k=1}^K &= \arg \min \left[C_{\text{NLS}} \left(\{\gamma(s_k), s_k\}_{k=1}^K \right) \right] \\ &= \arg \min \|\mathbf{g} - \mathbf{R}\gamma\|^2 \end{aligned} \quad (16)$$

where $\|\cdot\|$ represents the Euclidean norm and K is the number of scatterers inside an azimuth-range pixel. The cost

function is expressed as

$$C_{\text{NLS}} \left(\{ \gamma(s_k), s_k \}_{k=1}^K \right) = \| \mathbf{g} - \mathbf{R}\boldsymbol{\gamma} \|^2$$

$$= \sum_{n=1}^N \left[g_n - \sum_{k=1}^K \gamma(s_k) \exp(j2\pi\xi_n s_k) \right]^2. \quad (17)$$

In the aforementioned equation, the NLS cost function is a nonlinear function for the elevation of the scatterer s_k but a linear function for the complex scattering intensity $\gamma(s_k)$ of scatterers, so this is a separable NLS estimation problem.

RELAX uses the separability of two parameters to estimate. First, the linear least-squares estimate of $\boldsymbol{\gamma}$ can be obtained by minimizing the NLS cost function according to $\boldsymbol{\gamma}$.

$$\hat{\boldsymbol{\gamma}} = (\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H \mathbf{g}. \quad (18)$$

Then, the elevations of the scatterers $s = [s_1, \dots, s_K]^T$ can be estimated by minimizing the following cost function:

$$C_s \left(\{s_k\}_{k=1}^K \right) = \left\| \left[\mathbf{I} - \mathbf{R}(\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H \right] \mathbf{g} \right\|^2 \quad (19)$$

where \mathbf{I} is the identity matrix. The solution of (19) is a nonlinear multidimensional optimization problem, which requires a multidimensional search of the global extremum, and the calculation amount is quite large. Based on the signal rejection mechanism, the RELAX algorithm simplifies complex multidimensional search problems into simpler multiple 1-D search problems by iterative methods.

B. RM-RELAX Method

As shown in Fig. 9, RM-RELAX needs to introduce the multilooking method and s_{ref} of each pixel into RELAX to achieve tomographic inversion. The principle and implementation of RM-RELAX are described in detail as follows.

1) *Pixel Grouping*: Based on the contour lines and the masks of individual buildings, pixels closest to the same contour line are grouped together to obtain n pixel groups, denoted as $\mathbf{G}_{i=1, \dots, n}$. This procedure is performed in the following three steps.

First, the contour lines can be translated toward both ends of the building mask with pixel step size to obtain many contour lines. An example is shown in Fig. 17(a). Second, the distances (in pixels along the range direction) between each pixel and its adjacent contour lines will be calculated. Finally, each pixel will be assigned to the closest contour line.

2) *Extend RELAX to M-RELAX*: The multilooking methods DCS and MCS have been demonstrated to perform better than the standard CS when dealing with a small data stack using simulated data [10]. Zhu *et al.* demonstrated the potential of DCS in real SAR data [11]. However, both DCS and MCS are based on the CS, which requires complexity calculation. RELAX is a slightly worse method than the CS in terms of superresolution, and has a reconstruction accuracy comparable to that of the CS. However, the computational complexity of RELAX is much lower than that of the CS. Therefore, we propose an M-RELAX method by combining a multilooking method and RELAX.

In the M-RELAX method, the basic formulas of the standard RELAX-based TomoSAR are used. However, the measurement

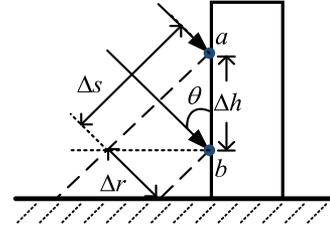


Fig. 10. Geometry of SAR imaging.

vector is extended by piling the independent pixel stacks, i.e.,

$$\mathbf{g}_{ML} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_M \end{bmatrix}. \quad (20)$$

The steering matrix must be concatenated accordingly, i.e.,

$$\mathbf{R}_{ML} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_M \end{bmatrix} \quad (21)$$

with $\mathbf{R}_1 = \mathbf{R}_2 = \dots = \mathbf{R}_M$. These M pixels can be extracted from the same pixel group \mathbf{G}_i .

3) *Calculate the Reference Elevation s_{ref} of Each Pixel*: GF-3 data are not only a small data stack but also have large a baseline interval. The large baseline interval will cause a severe ambiguity issue in elevation imaging. In [28], the range of elevation ambiguity was introduced as

$$S_{\text{amb}} = \lambda r_0 / 2\Delta b \quad (22)$$

where Δb is the interval of the orthogonal baseline. The Δb of real data is nonuniform, so we introduce a concept here called minimum range of elevation ambiguity (MREA), which can be expressed as

$$S_{\text{amb}}^{\text{min}} = \lambda r_0 / 2\Delta b_{\text{max}} \quad (23)$$

where Δb_{max} represents the maximum interval of orthogonal baselines. In [27], the author proved that in the case of nonuniform sampling, no aliasing target is generated within MREA. Therefore, we can reduce the interference of strong aliasing targets and improve the accuracy of the 3-D reconstruction by limiting ΔS (the discrete sampling range of the elevation direction) within an MREA for each pixel. We need to find a reference elevation s_{ref} for each pixel as the center of ΔS .

In this article, the contour lines and the SAR imaging geometry of high-rise buildings (shown in Fig. 10) can be combined to estimate the reference elevation s_{ref} of each pixel. The relationships of the variables in Fig. 10 can be expressed as

$$\Delta s = \Delta h / \sin \theta$$

$$\Delta h = \Delta r / \cos \theta \quad (24)$$

where θ is the incident angle, Δh is the height distance of two points (two pixels in the SAR image) a and b , and Δr and Δs are the distances of the two points along the range direction and elevation direction, respectively. We assume that the elevation of point b is s_b , and the elevation of point a can be calculated as

$$\begin{aligned} s_a &= s_b + \Delta s \\ &= s_b + \Delta r / (\sin \theta \cdot \cos \theta). \end{aligned} \quad (25)$$

The value of Δr can be obtained from the data parameter file. Therefore, the relative distance Δs is trustworthy and precise. If s_b is accurate, then s_a can also be considered accurate.

Actually, a pixel group will be selected from $\mathbf{G}_{i=1,\dots,n}$ as the reference pixel group \mathbf{G}_{ref} , and its elevation is 0. Then, the reference elevations $s_{\text{ref}i=1,\dots,n}$ of $\mathbf{G}_{i=1,\dots,n}$ relative to \mathbf{G}_{ref} can be calculated by (25). For each pixel, we determine the pixel group number i of the pixel first, and then, its reference elevation is $s_{\text{ref}i}$.

4) *Take s_{ref} to M-RELAX to Implement RM-RELAX:* After obtaining the reference elevation s_{ref} of each pixel, the ΔS of tomographic imaging can be limited as

$$s_{\text{ref}} - \frac{S_{\text{amb}}^{\text{min}}}{2} \leq \Delta S \leq s_{\text{ref}} + \frac{S_{\text{amb}}^{\text{min}}}{2}. \quad (26)$$

The new ΔS is substituted into M-RELAX, and the steering matrix \mathbf{R}_{ML} is rewritten as \mathbf{R}_{RML} , i.e.,

$$\mathbf{R}_{\text{RML}} = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{R}_2 \\ \vdots \\ \mathbf{R}_M \end{bmatrix} \quad (27)$$

where

$$\begin{aligned} \mathbf{R}_1 &= \mathbf{R}_2 = \dots = \mathbf{R}_M \\ &= \begin{bmatrix} \exp(j2\pi\xi_1 s_1) & \dots & \exp(j2\pi\xi_1 s_L) \\ \vdots & & \\ \exp(j2\pi\xi_N s_1) & \dots & \exp(j2\pi\xi_N s_L) \end{bmatrix}_{N \times L} \end{aligned} \quad (28)$$

where $s_1 = s_{\text{ref}} - S_{\text{amb}}^{\text{min}}/2$ and $s_L = s_{\text{ref}} + S_{\text{amb}}^{\text{min}}/2$. Since s_{ref} is a relative elevation, the measurement vector \mathbf{g}_{ML} should be rewritten as \mathbf{g}_{RML} , i.e.,

$$\mathbf{g}_{\text{RML}} = \begin{bmatrix} \mathbf{g}_1 \odot \mathbf{g}_{1_{\text{ref}}}^* \\ \mathbf{g}_2 \odot \mathbf{g}_{2_{\text{ref}}}^* \\ \vdots \\ \mathbf{g}_M \odot \mathbf{g}_{M_{\text{ref}}}^* \end{bmatrix} \quad (29)$$

where \odot represents the Hadamard product, $(\cdot)^*$ denotes the conjugate operation, and $\mathbf{g}_{m_{\text{ref}}} \in \mathbf{G}_{\text{ref}}$, $m = [1, 2, \dots, M]$ is the measurement vector of the reference pixel. Then, \mathbf{g}_{RML} and \mathbf{R}_{RML} can be taken into (18) and (19) to estimate the parameters of scatterers inside one pixel. Since we take the reference elevation into M-RELAX, we can call this method reference-elevation M-RELAX (RM-RELAX).

It should be noted that when using RM-RELAX for parameter estimation, it is necessary to assume the number K of scatterers first. For each pixel, we assume the number of scatterers is $K = 4$, and then, RM-RELAX can be used to estimate the parameters $\{\hat{\gamma}(s_k), s_k\}_{k=1}^K$.

5) *Model Selection:* In the presence of noise, the initial estimates $\{\hat{\gamma}(s_k), s_k\}_{k=1}^K$ may contain artifacts. To exactly detect the number of scatterers from the estimates, model selection should be executed after tomographic imaging [3], [23]. Let $\hat{\theta}(K)$ represent the estimates $\{\hat{\gamma}(s_k), s_k\}_{k=1}^K$ of the K scatterers, acquired by RM-RELAX. The model selection can be regarded as the following optimization problem:

$$\hat{K} = \arg \min_K \left\{ -2 \ln p(\mathbf{g} | \hat{\theta}(K), K) + 2C(K) \right\} \quad (30)$$

where $p(\mathbf{g} | \hat{\theta}(K), K)$ denotes the likelihood function of \mathbf{g} and $\hat{\theta}(K)$. Under the assumption that the model errors or disturbances $\mathbf{e} = \mathbf{g} - \mathbf{R}\hat{\gamma}(\hat{\theta}(K))$ are circular Gaussian distributed with zero mean and a covariance matrix of $\mathbf{C}_{\varepsilon\varepsilon} = |\varepsilon|^2 \mathbf{I}$, the likelihood function can be written as [3]

$$p(\mathbf{g} | \hat{\theta}(K), K) = \frac{1}{\pi^N |\varepsilon|^{2N}} \exp \left(-\frac{1}{|\varepsilon|^2} \left\| \mathbf{g} - \mathbf{R}\hat{\gamma}(\hat{\theta}(K)) \right\|^2 \right) \quad (31)$$

where $\hat{\gamma}(\hat{\theta}(K))$ is the reflectivity profile, and ε can be estimated by using the singular-value decomposition (SVD) method, which is introduced in [3]. The $C(K)$ in (30) is the penalty term to avoid overfitting of the data. Many types of penalty terms have been proposed, e.g., the Akaike information criterion (AIC), Bayesian information criterion (BIC) [29], and minimum description length [30]. In this article, we use the BIC penalized likelihood criterion with a penalty term $C(K) = 1.5K \ln N$ for model selection. The model \hat{K} is the one with the lowest penalized likelihood criterion value according to (30). Then, the parameters of scatterers inside a pixel can be determined as $\{\hat{\gamma}(s_k), s_k\}_{k=1}^{\hat{K}}$. The procedure of the high-rise building 3-D reconstruction with the RM-RELAX method is summarized in Algorithm 1.

C. CRLB for Elevation Estimation

The CRLB of elevation estimates for the single scatterer case is calculated according to (3). For the case of high-rise buildings, multiple scatterers are often mapped onto one azimuth-range pixel. The closely spaced scatterers will interfere with each other, rendering a degraded estimation accuracy of individual scatterers, which is described in detail in [23]. The CRLB for the elevation estimate of the q th ($q = 1, 2$ in this article) scatterer is introduced in [23] as

$$\sigma_{s_q} = c_0 \cdot \sigma_s \quad (32)$$

where c_0 is the interference factor depending on the elevation distance of two scatterers, i.e.,

$$c_0 \approx \max \left\{ \sqrt{2.57(\alpha^{-1.5} - 0.11) + 0.62}, 1 \right\} \quad (33)$$

Algorithm 1: The Procedure of the High-Rise Building 3-D Reconstruction With the RM-RELAX Method.

Require: The contour lines and the building mask \mathbf{B}_{mask} of a high-rise building.

- 1: Pixels are grouped as $\mathbf{G}_{i=1,\dots,n}$, and one group is selected as \mathbf{G}_{ref} with $s_{\text{ref}} = 0$;
 - 2: $s_{\text{ref}_{i=1,\dots,n}} \leftarrow$: calculate elevations of $\mathbf{G}_{i=1,\dots,n}$ relative to \mathbf{G}_{ref} by using (24) and (25);
 - 3: **for** each pixel $p \in \mathbf{B}_{\text{mask}}$ **do**
 - 4: Determine the group number i of this pixel, i.e., $p \in \mathbf{G}_i$;
 - 5: $\mathbf{g}_{\text{ML}} \leftarrow$: $\mathbf{g}_{m=1,\dots,M} \in \mathbf{G}_i$, $\mathbf{g}_{\text{ref}} \leftarrow$: $\mathbf{g}_{m=1,\dots,M} \in \mathbf{G}_{\text{ref}}$, $\Delta S \leftarrow$: take s_{ref_i} into (26);
 - 6: $\mathbf{g}_{\text{RML}} \leftarrow$: take \mathbf{g}_{ML} and \mathbf{g}_{ref} into (29), $\mathbf{R}_{\text{RML}} \leftarrow$: take ΔS into (27);
 - 7: $\{\gamma(s_{p_k}), s_{p_k}\}_{k=1}^K \leftarrow$: take \mathbf{g}_{RML} and \mathbf{R}_{RML} into (18) and (19);
 - 8: $\hat{K} \leftarrow$: take \mathbf{g}_{RML} and $\{\gamma(s_{p_k}), s_{p_k}\}_{k=1}^K$ into (30), $\{\gamma(s_{p_k}), s_{p_k}\}_{k=1}^{\hat{K}} \leftarrow$: $\{\gamma(s_{p_k}), s_{p_k}\}_{k=1}^K$;
 - 9: **end for**
-

where $\alpha = \Delta s / \rho_s$ is the elevation distance Δs between two scatterers normalized w.r.t. the Rayleigh resolution unit [23]. The value of c_0 is equal to one when the elevation distance of two scatterers is long, i.e., $\alpha \gg 1$, which means no interference between two scatterers. As Δs decreases, the value of c_0 increases, which means that the interference of the two scatterers is more serious.

V. EXPERIMENTAL RESULTS

In this section, we perform our proposed workflow, which includes the CLE and RM-RELAX methods, on simulated data and GF-3 data.

A. Performance Evaluation Using Simulated Data

In this subsection, the performance of the proposed RM-RELAX method, including the elevation estimation accuracy of single scatterers and the detection rate of separating overlaid scatterers, is evaluated using simulated data. Because RM-RELAX utilizes more measurements than RELAX and the priori knowledge about the reference elevation (available by using contour lines), it is natural for us to expect it to perform better than RELAX. We will also compare and evaluate the performance of the MCS and the proposed RM-RELAX method.

The parameters of the simulated data use the parameters of GF-3 as shown in Table I, and the baseline distribution is shown in Fig. 2(a). We use different methods, including RELAX, MCS, and the proposed RM-RELAX, to perform single- and double-scatterer experiments. In these experiments, we independently generate $M = 11$ snapshots for multilooking methods MCS and RM-RELAX and randomly extract one snapshot for RELAX. The reference elevation s_{ref} is calculated by

$$s_{\text{ref}} = s_f + s_e \quad (34)$$

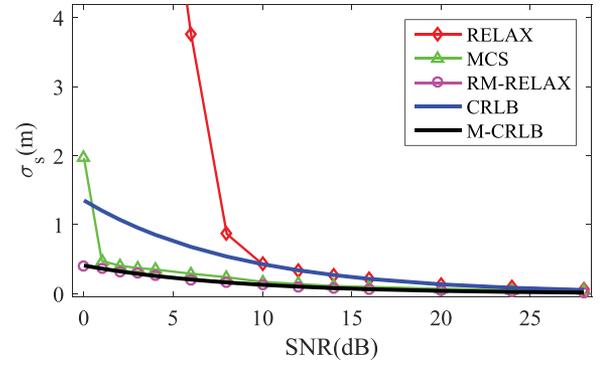


Fig. 11. Simulated results of single scatterers. Note that the M-CRLB is CRLB with $N \times M$ baselines.

where s_f is the true elevation of facade scatterers and s_e is the estimation error of s_f , which is a random value uniformly distributed between -4 and 4 m in experiments. Each result in the experiments is obtained after 2000 Monte Carlo experiments.

In the experiment of single scatterers, we set the elevation to 30 m. Fig. 11 shows the elevation estimated accuracy σ_s (root mean square error, RMSE) of RELAX, MCS and the proposed RM-RELAX compared to the CRLB under different SNR levels (the M-CRLB is CRLB with $N \times M$ baselines) [11], [31]. The experimental results can be analyzed as follows.

- 1) RELAX performs closely to CRLB at high SNR levels; however, the accuracy is severely degraded when the SNR is lower than 8 dB. It tells us that the single-look method RELAX is not suitable for 3-D inversion of single scatterers with SNR lower than 8 dB using small data stacks.
- 2) The performance of the MCS is much better than that of RELAX. The MCS uses the multilooking method for better antinoise performance. However, when the SNR is close to 0 dB, the accuracy of elevation estimation degrades suddenly. In the results of 2000 Monte Carlo experiments, we can find a few results with large errors in elevation estimation due to aliasing targets outside the MREA.
- 3) The RM-RELAX proposed in this article performs best. Its RMSE is always close to M-CRLB, and even the SNR drops to 0 dB. Because of the limitation of ΔS , RM-RELAX avoids detecting aliasing targets outside the MREA and obtains a more precise estimation.

In the experiment of double scatterers, we simulate the facade-ground interaction of two scatterers spaced by increasing elevation distance, which is a common model for urban high-rise buildings. Due to the interference factor depending on the distance of two scatterers (introduced in Section IV-C), we set the range of elevation distance Δs from 0 to 20 m, i.e., the ground-level scatterers are fixed at 0 m, and elevation of the facade-level scatterers varies from 0 to 20 m. The value of the SNR is 3 and 10 in dB. The simulated results are shown in Figs. 12 and 13. In advance, there are some legends to explain about the figure. The two solid line segments in the subplots of

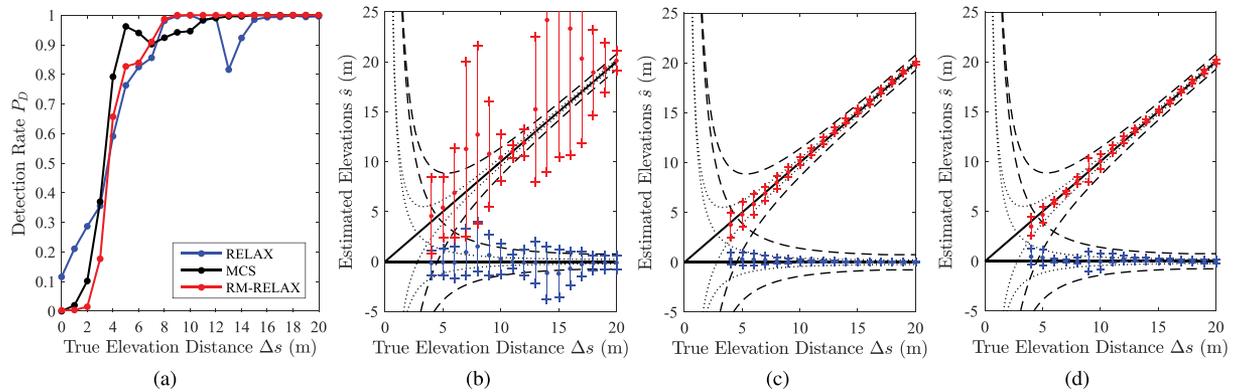


Fig. 12. Elevation estimated results of facade and ground scatterers with SNR = 10 dB. (a) Detection rate P_D and the estimated results using the methods of (b) RELAX, (c) MCS, and (d) RM-RELAX, with different elevation distances Δs .

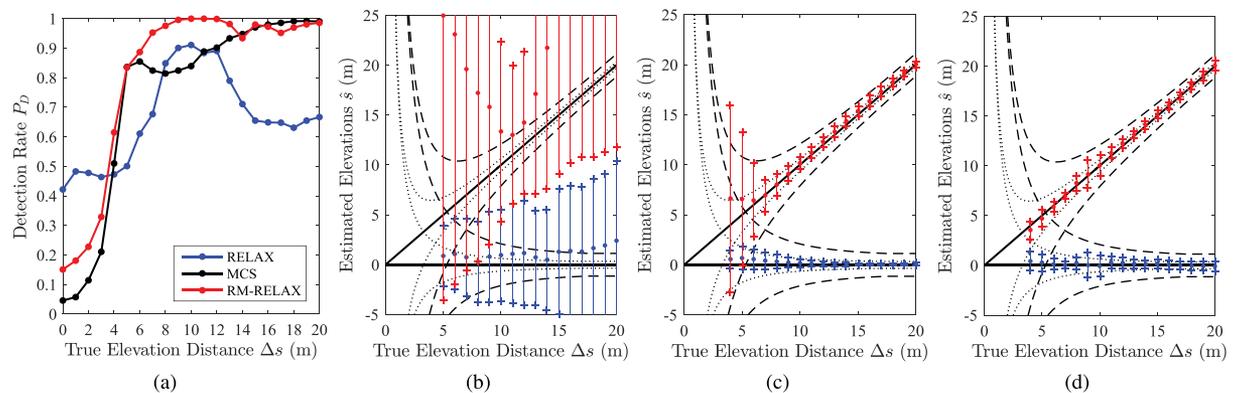


Fig. 13. Elevation estimated results of facade and ground scatterers with SNR = 3 dB. (a) Detection rate P_D and the estimated results using the methods of (b) RELAX, (c) MCS, and (d) RM-RELAX, with different elevation distances Δs .

Fig. 12 [except subplot (a)] are the true elevation of the simulated facade and ground scatterers, the dashed lines mark the true elevation $\pm 1 \times \text{CRLB}$ with N , denoted as $\text{CRLB}(N)$, and the dotted lines mark the true elevation $\pm 1 \times \text{CRLB}$ with NM , denoted as $\text{CRLB}(NM)$.

Figs. 12 and 13 show the simulated results with the SNR = 10 dB and SNR = 3 dB. Subplot (a) shows the detection rate P_D of different Δs , and (b)–(d) show the elevation estimated results of the facade and ground scatterers using the RELAX, MCS, and RM-RELAX methods. Each dot colored with red and blue depicts the mean value of all estimates, with the error bar indicating its standard deviation. The missing dots suggest that the detection rate is below 50%. Note that we define detection for the case where two scatterers are separated. For SNR = 10 dB, the detection rates of these three methods are similar, but the elevation estimates of RELAX are much worse than the estimates of MCS and RM-RELAX. Fig. 12(b) shows that most dots are far from the true elevation segments, and the error bars also exceed the scope of the $\text{CRLB}(N)$. The elevations reconstructed using MCS and RM-RELAX are reasonable and comparable. The dots and error bars always follow closely with the true elevation segments and the dotted lines $\text{CRLB}(NM)$,

respectively. For SNR = 3 dB, the detection rate of RELAX fluctuates. Combined with the estimated results of RELAX shown in Fig. 13(b), we can conclude that RELAX is very unreliable for elevation reconstruction with low SNR. The results of the MCS are reasonable when Δs is greater than 7 m, and the estimates of facade scatterers begin to deteriorate when Δs is less than 7 m due to the appearance of some outliers. Thanks to the reference elevation and limited ΔS of facade scatterers, the estimates of RM-RELAX are always reasonable, with Δs varying between 4 and 20 m.

From the simulated results of single and double scatterers, some conclusions can be drawn as follows.

- 1) In the case of single scatterers, the estimated accuracy of the proposed RM-RELAX is much better than that of RELAX when the SNR is below 10 dB and better than that of the MCS when the SNR drops to 0 dB.
- 2) In the case of double scatterers, the estimated accuracy of the RM-RELAX is always reasonable when Δs changes from 4 to 20 m and much better than that of RELAX. When the SNR is high (10 dB), the results of MCS and RM-RELAX are reasonable and comparable. However, when the SNR is low (3 dB), RM-RELAX performs better than

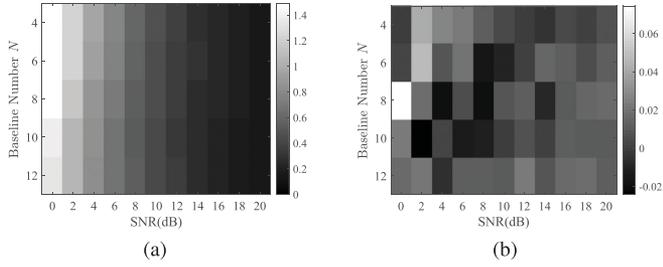


Fig. 14. (a) Elevation estimated accuracy of single scatterers. (b) Difference between the estimated accuracy and M-CRLB.

the MCS with Δs below 7 m because it avoids detecting outliers by using the reference elevation and limited ΔS of the facade scatterers.

To further demonstrate the applicability of the proposed RM-RELAX method, we also evaluate the performance of RM-RELAX with respect to the number of baselines. The system parameters of TerraSAR-X are adopted to make the experiment more general, where $\lambda = 0.311$ m, $r_0 = 621$ km, and the span of baselines B is approximately 300 m. We vary the number of baselines $N = 4 : 2 : 12$, and the baseline distribution is $\mathbf{b} = \{-149, \{a_i \cdot 300\}_{i=1}^{N-2}, 135\}$ m, where a_i is a uniformly distributed random number in $[-0.5, 0.5]$. The looks are still 11 for RM-RELAX. The experimental scheme still includes single and double scatterers. We set the elevation of single scatterers to 30 m, and the distance of double scatterers is the normalized elevation distance $\Delta\alpha = 0.2 : 0.1 : 1.2$ (α is introduced in Section IV-C). Figs. 14 and 15 show the simulated results of single and double scatterers, respectively. The performance of RM-RELAX can be summarized as follows.

- 1) For single scatterers, when the number of baselines is 4 or more, the elevation of single scatterers can be estimated well using the RM-RELAX method.
- 2) For double scatterers, the detection rate decreases as the baseline number N decreases. When the span of baselines is large ($B \approx 900$ m) and the number of baselines is small ($N = 4$), the detection rate will be severely reduced (as shown in the top two pictures in Fig. 15). The reason is that the increase in the baseline interval can lead to aggravation of the ambiguity issue, which will cause a large deviation in the estimation of ground scatterers (without prior knowledge on the reference elevation).
- 3) We can set the detection rate $P_D \geq 80\%$ of double scatterers with $\Delta\alpha = 1$ as the judgment threshold for the applicability of the method. From the aforementioned experiments, we recommend that the baseline number $N \geq 6$ for tomographic imaging using the proposed RM-RELAX method.

B. Practical Demonstration With GF-3 Data

The real data used in this experiment contain six SAR images acquired by China's spaceborne GF-3 satellite over Beijing,

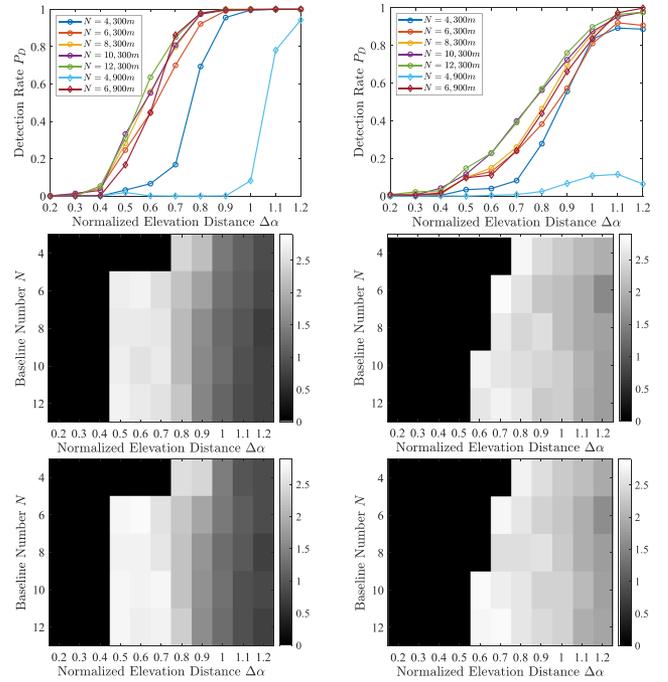


Fig. 15. Simulated results of double scatterers. From top to bottom: detection rate, RMSE (0 means $P_D \leq 20\%$) of ground and facade scatterers; from left to right: SNR = 10 and 3 dB.

China. The parameters of this data stack are shown in Table I. Fig. 1 shows the SAR image of the area of interest. The red box area is an interesting building in this experiment.

First, the proposed CLE method described in Section III is used to extract contour lines of interesting buildings as prior knowledge for subsequent processing. The results are shown in Fig. 16. The subplots (a) and (b) show the optical images of interesting buildings in Google Earth. The horizontal sections of these buildings are rectangular (Area 1) and oval-like (Area 2). The results of CLE shown in subplots (c) and (d) also conform to the shapes of these buildings. There may be small deviations (1 or 2 pixels), but the influence of the deviations is small or even negligible for subsequent MCS and RM-RELAX imaging. These results demonstrate that the proposed CLE method is suitable for high-rise buildings with simple shapes, which is consistent with most practical situations. We can notice that there are two buildings within the yellow elliptical area in subplot (a). However, because of the illumination angle of the SAR sensor, the two buildings are mixed together in the SAR image. It is no problem to separate the two buildings through advanced detection or classification methods, but it is not the purpose of this article, so the two buildings will be treated as one building in this article.

After extraction, contour lines can be used to obtain the pixel groups for MCS and RM-RELAX and calculate the reference elevation s_{ref} of each pixel for RM-RELAX. We select building A in the yellow box in Fig. 6 as an example to show the processing in detail. The steps of CLE of building A are shown

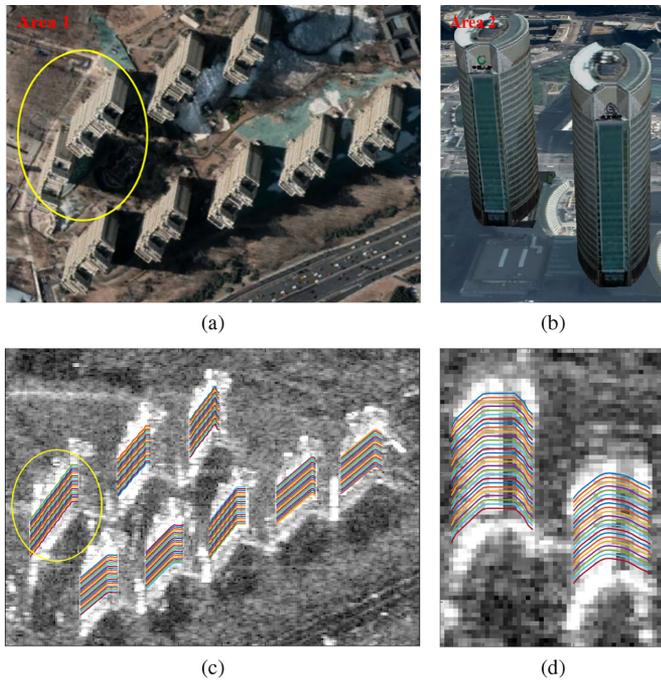


Fig. 16. Results of CLE. (a) and (b) Optical images of interesting buildings in Google Earth. (c) and (d) Color-coded contour lines of buildings in SAR images.

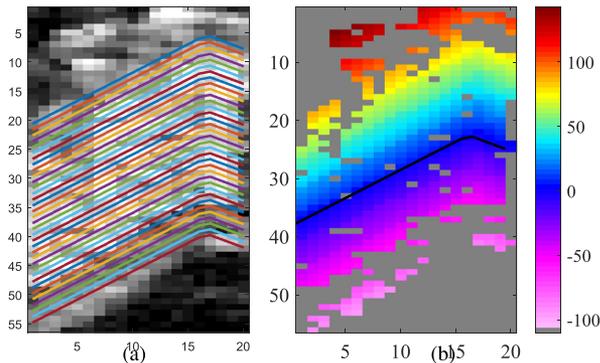


Fig. 17. (a) Color-coded contour lines and (b) each pixel's reference elevation of the building A.

in Section III. After extracting the contour line, we move the contour line up and down along the range direction of the SAR image to obtain many contour lines, which are color coded, as shown in Fig. 17(a). Then, each pixel will be grouped into the closest contour line. We take one pixel group as the reference group with $s_{ref} = 0$, which is shown as a black contour line in Fig. 17(b). The reference elevation s_{ref} of each pixel can be calculated by using (25) according to the reference group. The result is shown in Fig. 17(b). Finally, we take s_{ref} into RM-RELAX to reconstruct the height of the building A.

Fig. 18 shows the results of the 3-D reconstruction by using RELAX, MCS, and RM-RELAX. The number of looks for MCS and RM-RELAX is $M = 11$. The results of the MCS are much better than those of RELAX, but due to the elevation ambiguity

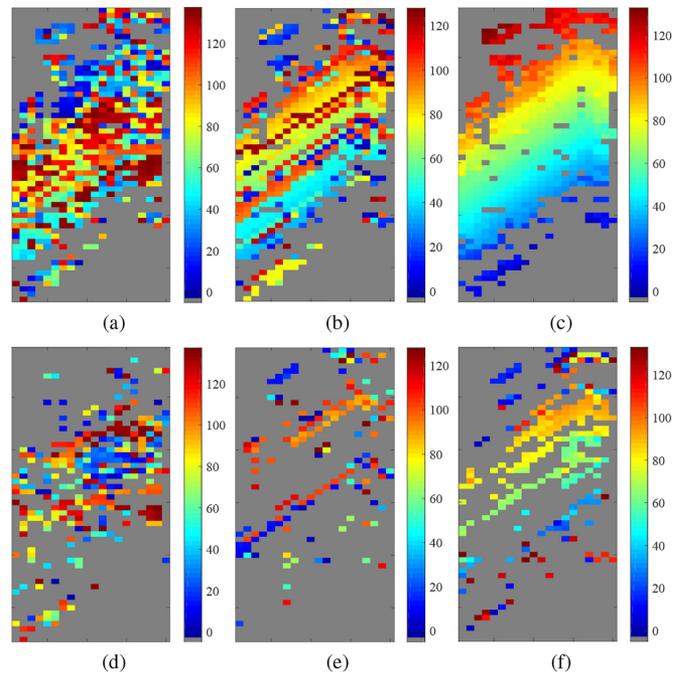


Fig. 18. Comparison of the 3-D reconstruction results of building A. (a) and (d) RELAX. (b) and (e) MCS. and (c) and (f) RM-RELAX. (a)-(c) are the first layer; (d)-(f) are the second layer.

issue, there are still some outliers. The RM-RELAX proposed in this article overcomes the ambiguity issue and obtains a more precise height estimation. To evaluate the accuracy of height estimation, we simulated a height model covering building A, which contains only the true height of the scatterers. The model decreases linearly from top to bottom with a maximum height of 133 m. Fig. 19 shows the comparison between the true height and the estimated height. Blue circles represent the first layer scatterers of building A, and red circles represent the second layer scatterers. As shown in Fig. 19, the estimated heights of most first layer scatterers show good agreement with the reference model. However, there are many outliers in the estimated results of RELAX and MCS. In contrast, the estimated heights of the first layer scatterers using the RM-RELAX method have no outliers. However, some suspected outliers appear in the estimated heights of the second layer scatterers on the top and bottom of the building A, which may be the second scatterer with lower amplitude or outliers caused by residual phase noise. We calculate the RMSEs of the estimated heights of the first layer scatterers with reference to the model as 57.8538, 40.0036, and 3.5278 m by using RELAX, MCS, and RM-RELAX, respectively. These results indicate that the estimates of RM-RELAX are better than those of RELAX and MCS.

Figs. 20 and 21 show the results of the 3-D reconstruction of the buildings in Area 1 and Area 2 in Fig. 1. These and the aforementioned results demonstrate the potential and applicability of the proposed workflow, including the CLE and RM-RELAX methods, for the 3-D reconstruction of high-rise buildings using a small data stack with large spatial baseline intervals.

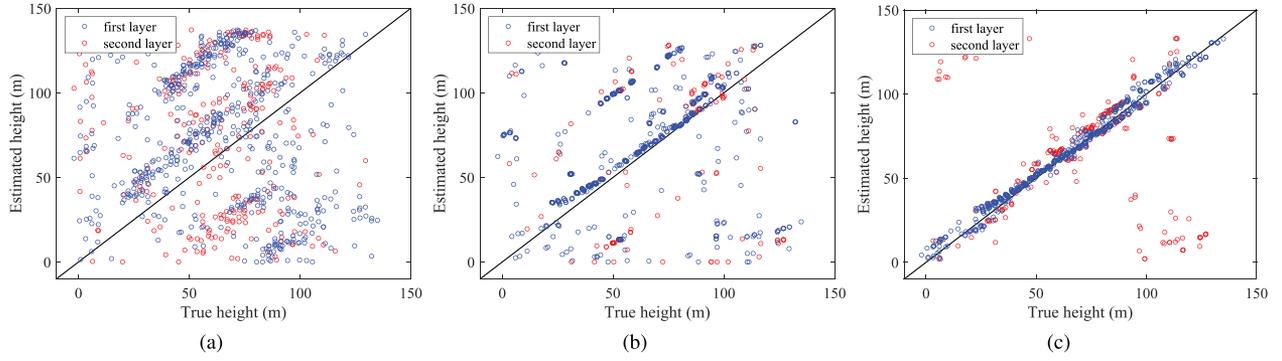


Fig. 19. Comparison of the true height and estimated height of building A. (a) RELAX. (b) MCS. (c) RM-RELAX.

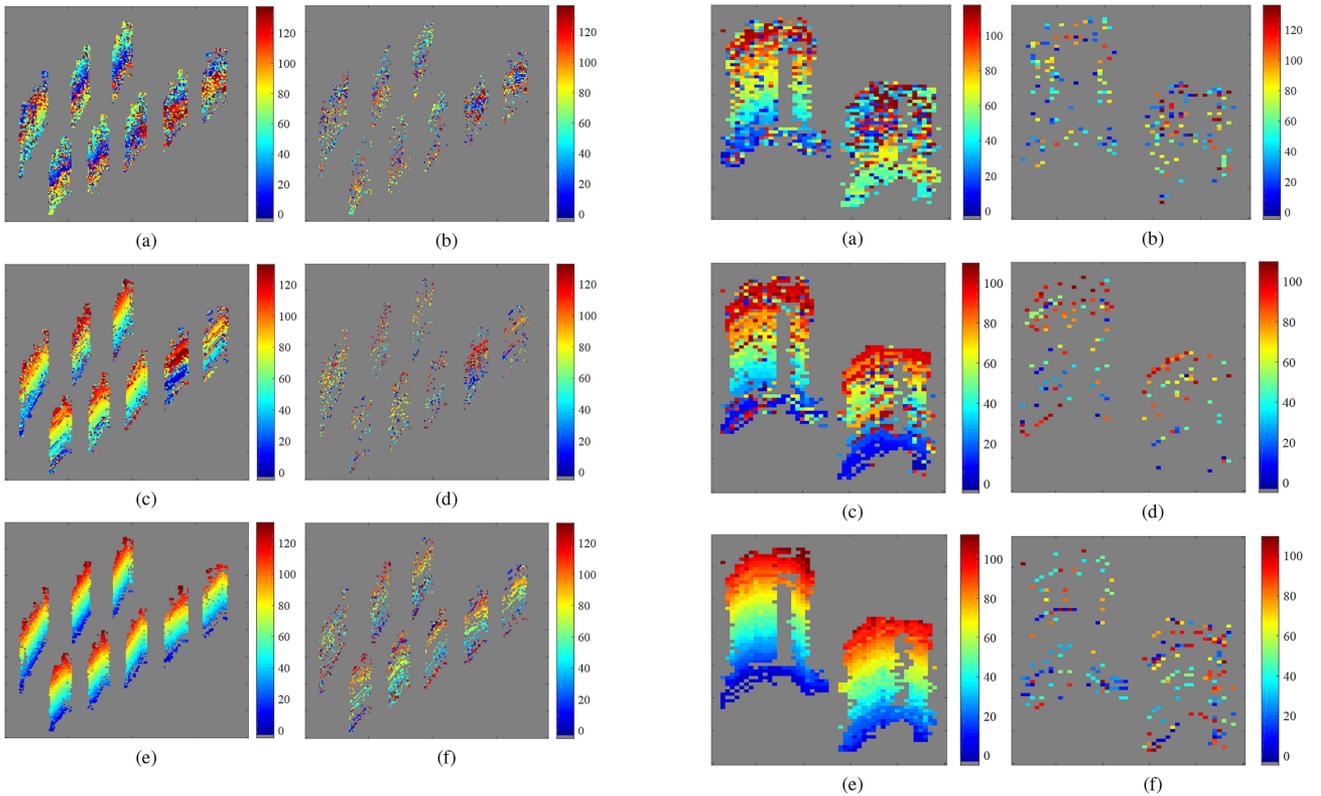


Fig. 20. Comparison of the 3-D reconstruction results of the Area 1. (a) and (b) RELAX. (c) and (d) MCS. (e) and (f) RM-RELAX. (a), (c), and (e) are the first layer; and (b), (d), and (f) are the second layer.

VI. CONCLUSION

In this article, a novel workflow is proposed, including the CLE and RM-RELAX methods, to obtain a more precise 3-D reconstruction of high-rise buildings with a small data stack. In this novel workflow, we introduced a data-driven CLE method to extract the contour lines without relying on any external data and extended the RELAX to RM-RELAX by using the idea of multilooking, reference elevation and MREA. We have demonstrated with simulated data and real data acquired by GF-3 satellite that the performance of RM-RELAX is better than that of other methods. The simulated results showed that the proposed RM-RELAX method can obtain more precise

Fig. 21. Comparison of the 3-D reconstruction results of the Area 2. (a) and (b) RELAX. (c) and (d) MCS. (e) and (f) RM-RELAX. (a), (c), and (e) are the first layer; and (b), (d), and (f) are the second layer.

elevation estimations than those of RELAX and MCS (in the case of low SNR). We also used the simulated experiments to evaluate the performance of RM-RELAX with respect to the number of baselines. Moreover, we used six SAR images with large baseline intervals to demonstrate the practicality and superiority of the proposed workflow for the 3-D reconstruction of high-rise buildings.

However, there are some limitations in the proposed workflow. The CLE method for extracting contour lines depends on the characteristics of the SAR data. It is suitable for relatively prominent or isolated high-rise buildings with general shapes and has difficulty dealing with buildings with complex shapes

and overlap. In the future, we will consider further improving and optimizing the proposed workflow to address more complex urban environments.

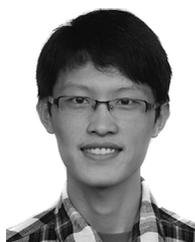
REFERENCES

- [1] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline 1-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.
- [2] 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.
- [3] 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.
- [4] P. Ma and H. Lin, "Robust detection of single and double persistent scatterers in urban built environments," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2124–2139, Apr. 2015.
- [5] 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.
- [6] 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, Nov. 2008.
- [7] F. Lombardini and A. Reigber, "Adaptive spectral estimation for multi-baseline SAR tomography with airborne 1-band data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2003, pp. 2014–2016.
- [8] F. Lombardini and F. Cai, "Temporal decorrelation-robust SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5412–5421, Dec. 2013.
- [9] X. Sun, A. Yu, Z. Dong, and D. Liang, "A high resolution imaging method for differential SAR tomography," *J. Electron. Inf. Technol.*, vol. 34, no. 2, pp. 273–278, Feb. 2012.
- [10] M. Schmitt and U. Stilla, "Compressive sensing based layover separation in airborne single-pass multi-baseline InSAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 2, pp. 313–317, Mar. 2012.
- [11] X. X. Zhu, N. Ge, and M. Shahzad, "Joint sparsity in SAR tomography for urban mapping," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 8, pp. 1498–1509, Dec. 2015.
- [12] Y. Shi, X. X. Zhu, and R. Bamler, "Nonlocal compressive sensing-based SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 5, pp. 3015–3024, Jan. 2019.
- [13] Y. Shi, Y. Wang, J. Kang, M. Lachaise, X. X. Zhu, and R. Bamler, "3D reconstruction from very small tandem-x stacks," in *Proc. 12th Eur. Conf. Synthetic Aperture Radar*, 2018, pp. 1–4.
- [14] Y. Wang, Y. Deng, R. Wang, and J. Wang, "Adaptive multilooking based on complex patch for multitemporal interferometry," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 3, pp. 907–918, Mar. 2018.
- [15] A. Paucullo, D. Reale, W. Franzé, and G. Fornaro, "Multi-look in GLRT-based detection of single and double persistent scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5125–5137, Sep. 2018.
- [16] S. Tebaldini, "Single and multipolarimetric SAR tomography of forested areas: A parametric approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2375–2387, May 2010.
- [17] Y. Huang, L. Ferro-Famil, and A. Reigber, "Under-foilage object imaging using SAR tomography and polarimetric spectral estimators," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2213–2225, Jun. 2012.
- [18] S. L. Wilson, "Two CFAR algorithms for interfering targets and nonhomogeneous clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 1, pp. 57–72, Jan. 1993.
- [19] G. Gao, L. Liu, L. Zhao, G. Shi, and G. Kuang, "An adaptive and fast CFAR algorithm based on automatic censoring for target detection in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 6, pp. 1685–1697, Jun. 2009.
- [20] L. Novak and S. Hesse, "On the performance of order-statistics CFAR detectors," in *Proc. IEEE 25th Asilomar Conf. Signals, Syst., Comput.*, 1991, vol. 2, pp. 835–840.
- [21] B. M. Kampes and N. Adam, "Velocity field retrieval from long term coherent points in radar interferometric stacks," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2003, pp. 941–943.
- [22] O. Mora, J. J. Mallorqui, and A. Broquetas, "Linear and nonlinear terrain deformation maps from a reduced set of interferometric SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 10, pp. 2243–2253, Oct. 2003.
- [23] 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.
- [24] A. Ferretti, C. Prati, and F. Rocca, "Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2202–2212, Sep. 2000.
- [25] J. Li and P. Stoica, "Efficient mixed-spectrum estimation with applications to target feature extraction," *IEEE Trans. Signal Process.*, vol. 44, no. 2, pp. 281–295, Feb. 1996.
- [26] Z.-S. Liu and J. Li, "Implementation of the relax algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 2, pp. 657–664, Apr. 1998.
- [27] X. Sun, "Research on SAR tomography and differential SAR tomography imaging technology," Ph.D. dissertation, Nat. Univ. Defense Technol., Changsha, China, 2012.
- [28] X. Sun, D. Zheng, Y. Anxi, and L. Diannong, "Super-resolution SAR tomographic imaging using ENVISAT-ASAR data," in *Proc. IEEE 3rd Int. Asia-Pacific Conf. Synthetic Aperture Radar*, 2011, pp. 1–4.
- [29] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding AIC and BIC in model selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, 2004.
- [30] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, no. 5, pp. 465–471, 1978.
- [31] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogramm. Fernerkundung Geoinf.*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.



Hongliang Lu (Student Member, IEEE) received the B.S. degree from Shandong University, Weihai, China, in 2016. He is currently working toward the Ph.D. degree with Space Microwave Remote Sensing System Department, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China.

He is currently with the University of Chinese Academy of Sciences, Beijing. His research interests include synthetic aperture radar tomography and signal processing.



Heng Zhang (Member, IEEE) received the B.S. degree from Nanjing University, Nanjing, China, in 2013, and the Ph.D. degree in communication and information systems from the University of Chinese Academy of Sciences, Beijing, China, in 2018.

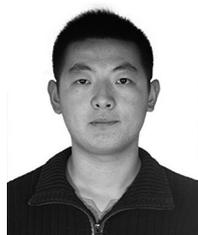
In 2018, he joined the Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing. His research interests include bistatic synthetic aperture radar imaging and interferometry.



Yunkai Deng (Member, IEEE) received the M.S. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 1993.

In 1993, he joined the Institute of Electronics, Chinese Academy of Sciences (IECAS), Beijing, where he was involved in antenna design, microwave circuit design, and spaceborne/airborne synthetic-aperture radar (SAR) technology. Since 1993, he has been a Research Fellow with the Department of Space Microwave Remote Sensing System, IECAS. He has been the Leader of several spaceborne/airborne SAR programs and developed some key technologies of spaceborne/airborne SAR. Since 2012, he has been a Principal Investigator with the Helmholtz-Chinese Academy of Sciences (CAS) Joint Research Group, Beijing, concerning Spaceborne Microwave Remote Sensing for Prevention and Forensic Analysis of Natural Hazards and Extreme Events. He is currently a Research Scientist with the University of Chinese Academy of Sciences, Beijing. He has authored or coauthored more than 100 articles since 2002, of which more than 100 peer-reviewed and well-known journal articles. His research interests include spaceborne/airborne SAR technology for advanced modes, multifunctional radar imaging, and microwave circuit design.

Mr. Deng is a Member of the Scientific Board. He was the recipient of several prizes, including the First and Second Class Rewards of National Defense Science and Technology Progress in 2007, the First Class Reward of the National Scientific and Technological Progress, in 2008, the achievements of the Outstanding Award of the CAS, in 2009, and the First Class Reward of Army Science and Technology Innovation, in 2016, for his outstanding contribution in SAR field.



Jili Wang received the B.S. degree from the Beijing Institute of Technology, Beijing, China, in 2013 and the Ph.D. degree in communication and information system from the University of Chinese Academy of Sciences, Beijing, in 2019. Since 2019, he joined the Spaceborne Microwave Remote Sensing System Department, Institute of Electronics, Chinese Academy of Sciences, Beijing.

From 2016 to 2017, he was a Research Assistant with the Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Shatin, Hong Kong. His research interests include developing multitemporal InSAR processing techniques and their applications on deformation parameter estimation.



Weidong Yu (Member, IEEE) was born in Henan, China, in 1969. He received the M.Sc. and Ph.D. degrees in electrical engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 1994 and 1997, respectively.

Since 1997, he has been with the Institute of Electronics, Chinese Academy of Sciences (IECAS), Beijing, China, where he became a Professor of Communication and Information Systems, in 2000. He has been the Chief Designer for several synthetic aperture radar (SAR) systems. He is currently the Director of Space Microwave Remote Sensing System Department, Aerospace Information Research Institute, Chinese Academy of Sciences. His research interests include spaceborne and airborne SAR system design and their signal processing.