URBAN 3D RECONSTRUCTION OF VHR SAR IMAGES USING ITERATIVE OPTIMIZATION ALGORITHM AND LAYOVER FIXED-ORDER MODEL

C. H. Zhang¹, L. Pang^{1*}, D.Y. Liu¹

¹School of Geomatics and Urban Spatial Informatics, Beijing University Of Civil Engineering And Architecture, No.15 Yongyuan Road, Daxing District, Beijing, China -210857002101@stu.bucea.edu.cn, panglei@bucea.edu.cn.

KEY WORDS: Iterative Optimization Algorithm Layeover Fixed-order Model TwIST-BIC, SAR Tomography (TomoSAR), 3D Reconstruction.

ABSTRACT:

Synthetic Aperture Radar (SAR) Tomography (TomoSAR) is a three-dimensional SAR imaging technique that uses multiple passes to process complex SAR images and obtain three-dimensional spatial scattering information to derive the elevation scattering distribution. Due to its own shortcomings, the elevation obtained by the traditional spectral estimation method has low resolution in the elevation direction and is affected by noise. The imaging algorithm based on compressed sensing can achieve super-resolution reconstruction in the elevation direction while reducing the number of observations. However, the CS algorithm still faces challenges when applied to real-world tomographic SAR imaging. In particular, it often requires numerous iterations to achieve satisfactory results, which significantly reduces its processing efficiency in large-scale tomography. To address the above issues, in this paper, we proposed an urban 3D reconstruction of VHR SAR images using an iterative optimization algorithm and layover fixed-order model. The iterative optimization algorithm and the layover fixed-order model consist of two parts: The TomoSAR imaging equation is solved by the two-step iterative shrinkage/thresholding (TwIST) algorithm, and the number of scatterers K is estimated by the Bayesian Information Criterion (BIC). In this paper, the effectiveness of TwIST-BIC in TomoSAR imaging in urban areas is verified with real TerraSARX data. By comparing with the OMP algorithm based on matching tracking and the FISTA algorithm based on gradient descent. The TWIST-BIC method is less complex, converges faster, and combines both execution speed and super-resolution, which can effectively solve the processing efficiency problem in large-area tomography and acquire high-resolution tomographic analysis.

1. INTRODUCTION

Unlike traditional optical observation methods, Synthetic Aperture Radar (SAR) has the ability to operate around the clock and in all weather conditions and is therefore widely used in areas such as land and resource surveying and natural disaster monitoring(Curlander & Mc Donough, 1991). However, conventional SAR imaging is limited to acquiring twodimensional images of the target in the azimuth-distance domain, which fails to accurately depict the three-dimensional scattering characteristics of the target. This limitation somewhat hinders the further application of SAR images. Synthetic Aperture Radar Tomography (TomoSAR) is an advanced extension of SAR imaging technology, enabling the acquisition of threedimensional spatial scattering information by processing SAR complex images obtained from multiple passes in the tomographic direction to derive the scattering distribution(Zhu et al., 2018a).

Satellite-based TomoSAR technology has been under development for about 20 years. In 2000, Reigber et al. performed a three-dimensional imaging experiment using a Fourier transform algorithm on 14-view L-band synthetic aperture radar (SAR) two-dimensional images. The results, obtained by imaging building structures, demonstrated for the first time the practical feasibility of SAR tomography imaging (Reigber & Moreira, 2000). In 2003, Fornaro et al. performed data processing experiments using ERS satellite-based SAR data to demonstrate the viability of using satellite-based SAR data for TomoSAR 3D imaging, laying the groundwork for subsequent research and practical applications (Fornaro et al., 2005). In 2010, Zhu et al. conducted the first research experiment on TomoSAR imaging of urban areas using compressed sensing techniques and proposed a minimized L1 parametric regularised TomoSAR imaging algorithm based on compressed sensing (Zhu & Bamler, 2010). In 2017, Wang et al. proposed an iterative reweighted alternating direction multiplier algorithm for fast TomoSAR imaging(Wang et al., 2017). In recent years, the field of synthetic aperture radar (SAR) imaging has witnessed a surge in novel imaging techniques incorporating deep learning methods, leading to significant research breakthroughs. The earliest attempt to extend traditional networks dates back to 2010, when Gregor and Le Cun proposed a novel network for fast sparse coding, LISTA, by evolving the Iterative Shrinkage and Thresholding Algorithm (ISTA)(Gregor & LeCun, 2010). In 2021, Qian Kun et al. applied the LISTA network to TomoSAR imaging and demonstrated through simulation experiments that the LISTA network could more accurately approximate the Cramér-Rao lower bound (CRLB) for tomographic estimation in scenarios with both single and multiple scatterers in the elevation direction(Qian et al., 2021). With the launch and deployment of a new class of meterresolution Synthetic Aperture Radar (SAR) systems (TerraSAR-X, COSMO-Skymed, and GaoFen-3), fast, high-resolution, multi-angle, layered three-dimensional imaging of large areas is expected in the future, further extending the range of TomoSAR applications.

In this paper, we proposed an experimental study on TomoSAR and imaging for high-quality 3D reconstruction of buildings using CS techniques based on 17-view TerraSAR-X SAR complex image data. The iterative optimization algorithms and layered fixed-order models are employed in the process. The

^{*} Corresponding author :panglei@bucea.edu.cn

subsequent structure of this paper is as follows: Section 2 focuses on the TomoSAR imaging model, Section 3 introduces the compressed sensing (CS) solution scheme for the TomoSAR model, and several technically promising methods for solving the TomoSAR equation are presented. In Section 4, a TomoSAR imaging study is conducted using SAR complex image data from 17-view TerraSAR-X, followed by a presentation and analysis of experimental results. Finally, Section 5 concludes the article.

2. TOMOSAR IMAGING MODEL

TomoSAR is achieved through the use of multiple heightoriented baselines, which can be achieved by performing multiple flight passes over time or by using multiple antennas simultaneously. The geometric model representing SAR Tomography imaging is shown in Figure 1.

Synthetic Aperture Radar (SAR) uses N observations of the same target, obtained from different spatial and temporal positions, to construct a synthetic aperture in the height direction. This allows three-dimensional imaging in terms of range-azimuth-elevation and facilitates the deconvolution of scatterers. The geometric model for tomographic SAR imaging is shown where r is the range direction, x is the azimuth direction and s is the elevation direction.



Fig. 1 Schematic diagram of TomoSAR imaging geometry

Assuming that N baselines of data (Fig. 1) are available, i.e., N+1 SAR images acquired by the antenna and arranged geometrically arranged, the pixel values at corresponding positions on these images can be formed into a sequence $y = [y_0, \dots, y_n]$ of length N+1 after amplitude and phase calibration. The N/2 image is selected as the primary image and each pixel value can be expressed as an integral of the scattering rate distribution along the upward slope. This can be expressed as:

$$y_n = \int_{\Delta s} \gamma(s) \exp(-j2\pi\zeta_n s) ds, n = 0, \dots, N$$
(1)

If the function is sampled discretely L times along the elevation direction and L is large enough, Eq. (1) can be discretised as:

$$\begin{pmatrix} g_1 \\ \vdots \\ g_N \end{pmatrix} = \begin{pmatrix} A_{11} & \cdots & A_{1L} \\ \vdots & \ddots & \vdots \\ A_{N1} & \cdots & A_{NL} \end{pmatrix} \cdot \begin{pmatrix} \gamma(s_1) \\ \vdots \\ \gamma(s_L) \end{pmatrix} + \varepsilon$$
(2)

where g is the measurement vector of length N, A is the dictionary matrix of size N×L, where $A_{n,l} = exp(-j2\pi\zeta_n s_l)$, N is the number of observations, $s_l(l = 1, 2, \dots, L)$ is the upward direction of elevation towards the sampling location, and ε is the noise vector.

After discretization, Eq. (2) can be simply approximated as follows(Fornaro et al., 2003):

$$g_{N\times 1} = A_{N\times L} \cdot \gamma_{L\times 1} + \varepsilon \tag{3}$$

The Tomographic resolution is related to the maximum vertical baseline length Δb (shown in Figure 1). Assuming that the baselines are uniformly distributed along the length of the synthetic aperture above the slant and are well sampled, the Rayleigh resolution above the slant-range direction can be obtained as Eq. (4), which can be calculated using the following:

$$\rho_s = \lambda \cdot r/2 \cdot \Delta b \tag{4}$$

where Δb is the baseline range, r is the slant-range distance from the sensor, and λ is the wavelength.

3. TOMOSAR INVERSION

In the absence of error effects, Spaceborne TomoSAR 3D imaging involves two essential steps: 3D imaging and scattering parameter extraction. Three-dimensional imaging involves reconstructing the complex scattering coefficients $\gamma(s)$ from the TomoSAR signals, while scattering parameter extraction involves determining the number and height position of each scattering element based on the reconstructed coefficients. This requires the use of scatterer detection methods (Ren et al., 2022). For the model described in Eq. (3), this paper implements compressed sensing based TomoSAR imaging by solving the following optimisation problem. The solution method can be divided into two main aspects: tomographic sparse spectrum estimation and model order selection.

3.1 Tomographic Spectrum Estimation

For 3D imaging using spaceborne SAR data, the spectral estimation method is commonly used to reconstruct the backscatter coefficients due to small variations in the angle of incidence between different observations(Jakowatz et al., 2012). Assuming that the elevation profile is sparse, a sparsity constraint (i.e., L0 norm minimisation) can be imposed by solving the underdetermined linear system of Eq. (3). The sparsest estimate of the reflectivity can be given by Eq. (5).

$$\min_{\gamma} \|\gamma\|_0 s. t. \|g - A\gamma\|_2^2 \le M\sigma_{\varepsilon}^2 \tag{5}$$

where $\|*\|_0$ is the L0 norm; M is the number of acquisitions;

 σ_{ε}^2 represents the model fitting error(Ren et al., 2022). However, the mathematical solution of the L0-norm minimization is a generalized NP-hard problem, whereas Cand ès et al. showed that the L0-norm minimization problem can be transformed into an L1-norm minimization problem if the signal X satisfies the conditions of incoherence, isometric constraints, etc., thus transforming the non-convex optimization model into a convex optimization model to be solved. Thus, the L0 norm can be approximated as the L1 norm. It can be approximated by ridge estimation as follows:

$$\stackrel{\wedge}{\gamma} = \arg\min_{\gamma} \left\{ \frac{1}{2} \|g - A\gamma\|_2^2 + \lambda_K \|\gamma\|_1 \right\}$$
(6)

where λ_K is a hyperparameter that needs to be adjusted according to the noise level. For Eq. (6), the current solution methods primarily consist of convex optimization algorithms and matching tracking class algorithms, which differ significantly in principle. The core of the convex optimization algorithm is to find the best matching atomic bases in each iteration to reconstruct the original signal with the highest accuracy using the

least number of bases. The match-tracking algorithm achieves a sparse approximation of the original signal by iteratively searching for the signal atom that best matches the measurement matrix. In Section 3.3 we present three iterative optimisation algorithms and matched tracing algorithms for solving the TomoSAR equation.

3.2 Model Order Selection

The task of model selection in SAR tomography is to select a matching statistical model from a set of possible parametric statistical models based on the estimated backscatter profile data. In estimates involving mixed distribution components and model order, the Bayesian Information Criterion(BIC) based on information theory(Schwarz, 1978) is the most commonly used estimation method. The likelihood criteria is given by:

$$\hat{K} = \arg\min_{K} \left\{ \frac{2\|g - A\gamma\|_{2}^{2}}{(M\sigma_{\varepsilon}^{2})} + 3KInM \right\}$$
(7)

Where $\frac{2\|g - A\gamma\|_2^2}{(M\sigma_{\varepsilon}^2)}$ is the log-likelihood term of the distribution model and is used to characterize the fit performance of the model, 3KInM is the model complexity penalty term; K is the number of scatterers in the stacked mask and M refers to the number of channels in the image. Since each scatterer corresponds to three parameters, amplitude, phase, and elevation, the model complexity corresponding to K scatterers is 3KInM.

3.3 OVERVIEW OF EXISTING ALGORITHMS

3.3.1 Fast Iterative Shrinkage Thresholding Algorithm (FISTA)

The Fast Iterative Shrinkage Threshold (FISTA) algorithm is a common method for solving L1 parametric minimization problems. It is mainly based on the principle of gradient descent for solving the convex optimization problem in Equation (6). The FISTA algorithm is based on the Iterative Shrinkage Threshold (ISTA) algorithm using the Nesterov acceleration technique, where the difference between one iteration and the current iteration is added as an acceleration term when the previous variable is updated, allowing the algorithm to converge quickly. The FISTA algorithm process is as follows:



Fig 2 a. Optical picture of the Study area (from Google Earth)

$$\gamma_{k} = T(z_{k})$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t^{2}}}{2}$$

$$z_{k+1} = \gamma_{k} + \frac{t_{k-1}}{t_{k+1}}(\gamma_{k} - \gamma_{k-1})$$
(8)

where, t_k is the iteration step it depends on the previous iteration step and is constantly changing. In FISTA algorithm z_k is a specific linear combination of γ_{k-1} and γ_{k-2} . By this specific linear combination the FISTA algorithm converges faster.

3.3.2 Two-step Iterative Shrinkage/Thresholding (TwIST)

The Two-Step Iterative Shrinkage/Thresholding (TwIST) method has been proposed to solve the pathological linear inverse problem. The algorithm is based on the Iterative Shrinkage Thresholding (IST) and the Iterative Reweighted Shrinkage (IRS) algorithms, and its core idea is to use the first two iteration values to update the current value. TomoSAR can be solved by minimizing the number of L1 paradigms. Therefore, the use of TwIST for the TomoSAR study is feasible(Wei et al., 2015). TwIST reconstructs the scattering coefficients by iteratively solving the objective function of Eq. (5). The initialisation of the TWIST method can be carried out by the following equation:

$$\gamma_1 = \psi \ (\gamma_0 + G^T (g - G\gamma_0)) \tag{9}$$

where γ_0 is the initial backward scattering quantity, which can be obtained by the least squares method, and ψ is the convex optimisation function defined in Eq. (6). When For $t \ge 1$, the iterative process starts until convergence to the minimum value of the objective function Eq. (6).

$$\gamma_{t+1} = (1 - \alpha)\gamma_{t-1} + (\alpha - \beta)\gamma_t + \beta \cdot \psi(\gamma_t + G^T(g - G\gamma_t))$$
(10)

where α and β are weighting parameters. The meaning of "twostep" here is that the value of γ_{t+1} depends on both γ_t and γ_{t-1} , not only on γ_t . This kind of the two-step structure enables fast and unbiased estimation of the pathological problem.



b. Cropped SAR image block (770x770 pixels)



Fig. 3. Spatial and temporal baseline of the data stack.

3.3.3 Orthogonal Matching Pursuit (OMP)

The orthogonal matching pursuit(OMP) algorithm is an improvement of the matching pursuit (MP) algorithm(Tsaig & Donoho, 2006), which compares the inner product of the target signal with the atoms in the selected measurement set, and selects the atoms that most closely match the measurement signal g from the sensing matrix each time, which are then included in the reconstructed sparse basis due to the orthogonalisation process, so as to enable the algorithm to converge quickly and obtain the reconstructed sparse basis that meets the accuracy of iterative error, and then reconstruct the original layer-dispersion vector. The OMP algorithm has been one of the commonly used algorithms in the CS field because of its simple structure and low computational complexity. However, the OMP algorithm is less robust to noise, and the reconstruction accuracy of the algorithm deteriorates when the atomic correlation in the measurement matrix is high.

4. PRACTICAL DATA PROCESSING AND ANALYSIS

4.1 Study area and datasets

For this study, a stack of 17 stripmap images was acquired by the TerraSAR-X sensor, enabling the evaluation of proposed TomoSAR methods. We selected the apartment building in Pingzhou Xincun, Baoan District, Shenzhen and Bao'an First

Stack Name	ShenZhen
Number of Images	17
Orbit	Ascending
Work Mode	Spotlight
Flight Height	541Km
Revisit Cycle	11 days
Polarization Method	HH
Azimuth Resolution	0.25m
Slant Range Resolution	0.8m
Baseline Range	189.8m
Expected resolution	20.953m

Table 1. Parameters of the TerraSAR-X data

Foreign Language School (Middle School) as the target area for TomoSAR imaging. As a result of the field survey, the tallest building in the study area is a 28-storey residential building with a height of approximately 84 metres. The image size of the area selected for the experiment is 770 x 770 pixels. Corresponding field area is approximately 0.48 x 0.5 Km2 The optical image is shown in Figure 2(a), and its corresponding SAR image area is shown in Figure 2(b).

TerraSAR-X is the world's first X-band Earth observation SAR satellite. As you can see from the XML file that accompanies the data: The satellite flew at an altitude of 514 kilometers, a revisit period of 11 days, an irradiation direction of right-side view, The polarization mode of the images is HH, and the values of the slant range and azimuth resolutions are 0.8 m and 0.25 m, respectively. The parameters of the TeraSAR-X SAR sensors are summarized in Table 1. The SAR images were acquired over Shenzhen, China, from the ascending orbit direction between 6 January 2016 and 31 July 2017. Since the image captured on 13 August 2016 was situated at the relative center of both temporal and spatial baselines, it was chosen as the reference master image. The vertical effective baseline of this SAR dataset is non-uniformly distributed, with a baseline span of approximately $b_{\perp} = 189.8m$ and a heightwise Rayleigh resolution of $\rho_s = 20.953m$.

The baseline was calculated using SARscape(Simonetto & Follin, 2011), and Figure 3 illustrates the temporal and spatial baselines of the dataset consisting of seventeen views.



Fig 4. TomoSAR processing framework.



Fig.5 Tomography results (after geocoding)

4.2 Experiments and Results

This section presents the results of the TomoSAR reconstruction based on the TerraSAR-X dataset. The image size of the selected area for the experiment is 770 x 770 pixels. The actual size of the corresponding area is approximately 0.48x0.50km². The processing flow of the data is shown in Figure 4 Firstly, in the pre-processing phase, we utilized the correlation function method for image alignment using SNAP, an open-source SAR processing toolbox developed under ESA's auspices. After coregistration, an interferometric image pair was generated and during its formation by SNAP, topographic phase removal was performed using an external DEM. Subsequently, operations such as flat earth phase and atmospheric phase screen removal were performed. Secondly, the SAR tomography processing was then performed using FISTA/TwIST/OMP, followed by Bayesian Information Criterion (BIC) to estimate the number of scatterers per pixel and their corresponding reflectance. Finally, project the estimated coordinates of the tomography scattering point locations from the SAR imaging coordinate system to the geographic coordinate system. This process can be realized by geocoding technique.

The distribution of the number of scatterers in the experimental area extracted using the three SAR tomography methods is shown in Figure 5, and we have assigned different colors to these scatterers according to the different elevation values of each scatterer. The accurate evaluation of this experiment comprises two components. Firstly, it involves comparing the detected estimation of the scatterer height values with the true height building to calculate the accuracy of height estimation for these scatterers. Secondly, the algorithm's performance is quantitatively assessed based on both the time required for the 3D reconstruction of the building and the density of the detected scatterers, as presented in Table 2. Figure 5 and Table 2 provide visual evidence supporting these findings.

Evaluation			
indicators	TwIST-BIC	FISTA-BIC	OMP-BIC
Number of			
scatterers	183552	166002	110597
Processing			
time	2403s	3762s	262s
Height			
estimates	84.35m	84.82	85.89

 Table 2. Algorithm reconfiguration performance comparison

 (Includes model order selection)

As depicted in Figure 5 and presented in Table 2. The TwIST algorithm provides better reconstruction completeness of the building as a whole compared to the OMP algorithm based on match tracing, with richer details on the edges and main body of the residential building, and with more structural integrity of the roof. TwIST, due to its unique "two-step" iteration mechanism, where the result of each iteration step depends on the results of the previous two steps, detected 183552 scatterers in a field area of 0.48x0.50 km², and the height of the scatterers was estimated to be 84.32 meters, which is very close to the height of the building, 84 meters. The OMP algorithm based on matched tracking indicates the best imaging efficiency, about ten times that of the TwIST algorithm, but due to the flaws of the algorithm itself, OMP has the worst reconstruction accuracy and the lowest number of detected scatterers. It is worth noting that the TwIST algorithm detects more than 20,000 more scatterers than the FISTA algorithm, which is attributed to the TwIST's unique "two-step iterative updating" mechanism, and is instead nearly

one-third faster than the FISTA algorithm in terms of processing time.

5. CONCLUSION

In this work, we proposed a 3D reconstruction of buildings using an iterative optimization algorithm and layover fixed-order method to improve the processing efficiency of algorithms based on the theory of compressed perception for real-world applications and to improve the accuracy of building height estimation in urban areas. The analysis of the proposed TomoSAR method is based on 17 TerraSAR-X strip map ascending orbit superimposed images of Bao'an District, Shenzhen. Therefore, the tallest building in Pengzhou Xincun district is selected as a case study building to analyze the efficiency and accuracy of the TwIST-BIC estimation method. The gradient descent-based FISTA algorithm and the matching tracking-based OMP algorithm are used as comparative experiments to evaluate the effectiveness of the method in terms of building reconstruction accuracy and imaging processing time. The results show that: The TwIST-BIC method is characterized by lower complexity, faster convergence, both execution speed and super-resolution capability. It is suitable for 3D imaging of large area scenes. It also demonstrates that the compressed perception technique has great potential for application in urban SAR tomography, particularly for high-resolution SAR tomography.

ACKNOWLEDGMENTS

The authors would like to thank Deutsches Zentrum für Luft-und Raumfahrt(DLR) for providing the SAR tomographic experiment datasets (TerraSAR-X proposal RES3668).

REFERENCES

- Curlander, J. C., & McDonough, R. N. (1991). Synthetic aperture radar (Vol. 11). Wiley, New York.
- Fornaro, G., Lombardini, F., & Serafino, F. (2005). Threedimensional multipass SAR focusing: Experiments with long-term spaceborne data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(4), 702-714.
- Fornaro, G., Serafino, F., & Soldovieri, F. (2003). Threedimensional focusing with multipass SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 41(3), 507-517.
- Gregor, K., & LeCun, Y. (2010). Learning fast approximations of sparse coding. Proceedings of the 27th international conference on international conference on machine learning,
- Jakowatz, C. V., Wahl, D. E., Eichel, P. H., Ghiglia, D. C., & Thompson, P. A. (2012). Spotlight-mode synthetic aperture radar: a signal processing approach: a signal processing approach. Springer Science & Business Media.
- Qian, K., Wang, Y., Shi, Y., & Zhu, X. X. (2021). Superresolving sar tomography using deep learning. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS,
- Reigber, A., & Moreira, A. (2000). First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38(5), 2142-2152.

- Ren, Y., Xiao, A., Hu, F., Xu, F., Qiu, X., Ding, C., & Jin, Y.-Q. (2022). Coprime sensing for airborne array interferometric SAR tomography. *IEEE Transactions* on Geoscience and Remote Sensing, 60, 1-15.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, 461-464.
- Simonetto, E., & Follin, J.-M. (2011). An overview on interferometric SAR software and a comparison between DORIS and SARSCAPE packages. Geospatial free and open source software in the 21st century: Proceedings of the first open source geospatial research symposium, OGRS 2009,
- Tsaig, Y., & Donoho, D. L. (2006). Extensions of compressed sensing. *Signal processing*, 86(3), 549-571.
- Wang, X., Xu, F., & Jin, Y.-Q. (2017). The iterative reweighted alternating direction method of multipliers for separating structural layovers in SAR tomography. *IEEE Geoscience and Remote Sensing Letters*, 14(11), 1883-1887.
- Wei, L., Balz, T., Zhang, L., & Liao, M. (2015). A novel fast approach for SAR tomography: Two-step iterative shrinkage/thresholding. *IEEE Geoscience and Remote Sensing Letters*, 12(6), 1377-1381.
- Zhu, X. X., & Bamler, R. (2010). Tomographic SAR inversion by \$ L_ {1} \$-norm regularization—The compressive sensing approach. *IEEE Transactions on Geoscience* and Remote Sensing, 48(10), 3839-3846.
- Zhu, X. X., Wang, Y., Montazeri, S., & Ge, N. (2018a). A Review of Ten-Year Advances of Multi-Baseline SAR Interferometry Using TerraSAR-X Data. *Remote* Sensing, 10(9), 1374. https://www.mdpi.com/2072-4292/10/9/1374