ADATOMO-NET: A NOVEL DEEP LEARNING APPROACH FOR SAR TOMOGRAPHY IMAGING AND AUTOFOCUSING

Yunqiao Hu, Xiaoling Zhang, Shunjun Wei, Yu Ren, Nan Wang, Jun Shi

School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, P. R. China, 611731

IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium | 978-1-6654-2792-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/IGARSS46834.2022.9884028

ABSTRACT

Tomographic Synthetic aperture radar (TomoSAR) imaging algorithms for urban areas based on Compressed Sensing (CS) often have high time complexity due to many times iterations. Moreover, the phase error (PE) that exists will defocus TomoSAR imaging results. To reduce PE in the TomoSAR process, researchers use methods such as PS-InSAR, phase gradient autofocusing (PGA), etc. However, these methods are computationally expensive, which hinders the application of fast high-resolution TomoSAR imaging. In this paper, we merge the PE compensation into FISTA framework and proposed a novel deep learning approach for TomoSAR imaging. The network is based on Ada-LFISTA architecture, dubbed as AdaTomo-Net. Experiment results show that AdaTomo-Net has higher imaging accuracy and considerable computational efficiency compared with typical CS algorithms and learning-based algorithms such as LISTA in the presence of PE.

***Index Terms*—** TomoSAR, Compressed sensing, Phase error, Phase gradient Autofocusing, Deep learning

1. INTRODUCTION

Tomographic Synthetic Aperture Radar (TomoSAR), as an advanced remote sensing technique, has the ability to provide 3D scattering profile of the observation area.It also has the capacity to explore 4D and even 5D information of the observation scene. Thus, TomoSAR has been widely used in urban areas monitoring, forest canopy height estimation, and other important remote sensing missions [1]. Compressed sensing (CS) algorithms are extensively used in TomoSAR imaging missions due to their super-resolution properties with insufficient number of observations. However, conventional CS-based TomoSAR imaging methods suffer from computationally expensive problem due to high computaionally complexity such as frequent matrix operations. When encountering large scale imaging scenes, this problem will become particularly severe.

In [2], the author proposed a deep learning way for 3D SAR imaging. They unfolded the VAMP algorithm to a deep neural network. The experiment results show that the deep network has the best reconstruction performance with superior efficiency in different SNRs, but the networks performance was only verified on the near-field measurement data in the microwave anechoic chamber, not on remote-sensing SAR data. In [3], the author applied LISTA to implement super-resolution TomoSAR imaging, and they verified the effectiveness of the proposed method on simulated data. However, the performance of LISTA relies on single steering matrix learning which makes it sensitive to the steering matrix perturbations [4]. When it comes to practical application, phase errors (PEs) always exist between the steering matrix and the real one due to the measurement errors of platform motions or atmospheric delay, which will make TomoSAR images defocused [5]. In recent years, many PE compensation methods have been proposed for super-resolution TomoSAR imaging as preprocessing methods such as PS-InSAR and Phase Gradient Autofocusing (PGA) [5]. However, these methods are computationally expensive, which limits their application on large scene TomoSAR imaging.

To address TomoSAR imaging with PE, in this paper, we unrolled the sparsity autofocus algorithm based on FISTA to a deep unfolded network for TomoSAR imaging dubbed as AdaTomo-Net, which combines PE compensation with the imaging process. The network is trained on simulated data sets and validated on simulated and real SAR data. Experiment results show that the proposed network has better imaging accuracy with high computational efficiency when the steering matrix has PE.

2. TOMOSAR MODEL WITH PHASE ERROR

In Single-Look pattern, the discrete TomoSAR model with PE can be expressed as [1]

$$y = \widetilde{\mathbf{A}}x + \epsilon \tag{1}$$

$$\widetilde{\mathbf{A}} = \mathbf{E}\mathbf{A} \tag{2}$$

where $y \in \mathbb{C}^{N\times 1}$ is the measured vector, $x \in \mathbb{C}^{M\times 1}$ is the scattering profile vector along the parpendicular to the line-of-sight (PLOS) direction, $\widetilde{\mathbf{A}}$ is the steering matrix with PE, $\mathbf{A} \in \mathbb{C}^{N\times M}$ is the steering matrix, ϵ denotes the noise term, and $\mathbf{E} \in \mathbb{C}^{N\times N}$ is the PE matrix which can be modeled as

$$\mathbf{E} = \mathrm{diag}\left\{e^{j\phi_1}, \ldots, e^{j\phi_N}\right\} \tag{3}$$

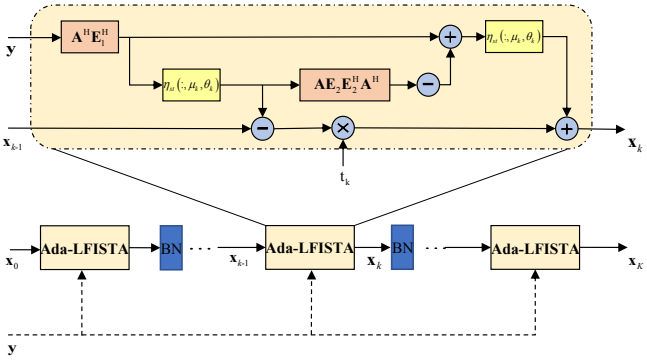


Fig. 1. AdaTomo-Net architecture: AdaTomo-Net is a K-step unrolled network, and the basic block of AdaTomo-Net is an Ada-LFISTA block, the output of the previous block passes through the Batch-Norm layer as input to the next block

3. TOMOSAR AUTOFOCUSING AND DEEP NETWORK IMPLEMENTATION

3.1. TomoSAR Autofocusing

Under the assumption of TomoSAR imaging, the optimal estimation of the scattering profile x can be obtained via solving a convex optimization with ℓ_1 constraint, which can be expressed as

$$\widehat{x} = \underset{x,\mathbf{E}}{\operatorname{argmin}} \|y - \mathbf{E}\mathbf{A}x\|_2^2 + \lambda \|x\|_1 \tag{4}$$

where λ is a regularization factor that controls the trade-off between data-fit term $\|y - \mathbf{E}\mathbf{A}x\|_2^2$ and sparsity term $\|x\|_1$. As (4) is a typical LASSO problem, there are many efficient solvers, e.g., ISTA, FISTA [6]. In this paper we choose FISTA as the foundation of our network for its convergence speed. In the canonical FISTA framework, the main iteration steps of TomoSAR autofocusing can be written as

$$x_{k+1} = \eta_{st}\left(v_k + \mu_k\left(\widehat{\mathbf{E}}_k\mathbf{A}\right)^{\mathrm{H}}\left(y - \widehat{\mathbf{E}}_k\mathbf{A}v_k\right), \theta_k\right) \tag{5}$$

$$\mathrm{t}_{k+1} = \frac{1+\sqrt{1+4\mathrm{t}_k^2}}{2} \tag{6}$$

$$v_{k+1} = x_{k+1} + \frac{\mathrm{t}_k - 1}{\mathrm{t}_{k+1}}\left(x_{k+1} - x_k\right) \tag{7}$$

where t_k is iteration step, and $\eta_{st}(\cdot)$ is the well known soft shrinkage function which can be defined as

$$\eta_{st}(r,\theta_k) = \operatorname{sign}(r)\max(|r| - \theta_k, 0) \tag{8}$$

According to [7], with the estimested value x_{k+1} fixed, the estimation of $\widehat{\mathbf{E}}_{k+1}$ of the kth step can be obtained by

$$\widehat{\mathbf{E}}_{k+1} = \operatorname{diag}\left(\exp\left(-j\angle\left(y^{\mathrm{H}}\mathbf{A}x_{k+1}\right)\right)\right) \tag{9}$$

3.2. Deep Network Implementation

According to our current knowledge, Ada-LFISTA, the variant of LISTA, is a good way to solve equation (1) due to its robustness to measurement matrix perturbations [8]. In addition, the architecture of the Ada-LFISTA fits well with the iteration steps (5)-(9). Therefore, the AdaTomo-Net implementation is based on the Ada-LFISTA. Fig 1 shows the basic architecture of AdaTomo-Net. $\mathbf{A}$ is the steering matrix that is embedded in the architecture. Different from the iteration step (9) listed in subsection 2.2, the phase error matrix $\widehat{\mathbf{E}}_k$ is replaced by learnable matrix $\mathbf{E}_1$, $\mathbf{E}_2$ and update during the network training process. The BatchNorm (BN) layer is used to prevent gradient explosion during training. TomoSAR data are in the complex domain, but Ada-LFISTA can not tackle complex data directly. To this end, we transfer TomoSAR data to the real domain in the same way in [9] and thus complex number operations in the network become real number operations.

4. EXPERIMENTS

4.1. Simulation Experiments

4.1.1. Simulation Setup

We refer to the system parameters of TerraSAR-X to set the simulation parameters, which are given in Table 1. The baseline configuration is also set according to the TerraSAR-X reference non-uniformly distributed baselines which are in the range -193m to 94m, and the number of acquisitions is 24, as is shown in Fig.2. The simulated SAR image of each observation is obtained by time-domain back-projection imaging algorithm[2].

Consider the lack of real measured TomoSAR data, we use simulated datasets for training. In the imaging range along the POLS direction of each range-azimuth pixel, we generate s randomly distributed point targets according to the sparsity of urban areas as a training label. Generally, $s \leq \frac{N}{2}$, and N denotes the observation number [9]. The point targets positions obey uniform distribution over the normalized elevation interval $[0,1]$, their amplitudes obey uniform distribution of $[0.5,1]$, and their phases obey uniform distribution of $[-\pi,\pi]$. In this way, we get 700,000 training labels, denoted as X. We add random perturbations to the steering matrix $\mathbf{A}$ according to (2). The perturbations are modeled as quadratic PE according to [10], in the form $\Phi_k = 0.5\pi\left((k-1)/N\right)^2$. Therefore, we get inputs Y for network training by adding different levels of noise, and the SNR levels are set to 3dB, 5dB, 10dB, 20dB, 30dB respectively. Then we choose the best training results.

4.1.2. Experiment with Simulated Point Targets

In the first simulation experiment, we randomly generate 10,000 test samples at different SNR levels with SNR = 0dB, 3dB, 5dB, 7dB, 10dB. Each test sample contains single

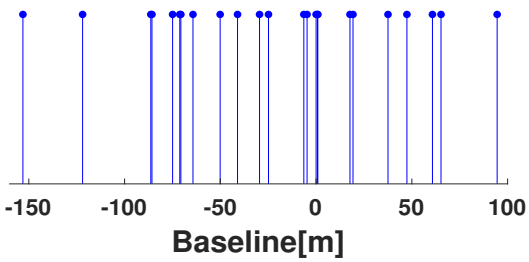


Fig. 2. Baseline Configuration.

Table 1. Simulation System Parameters

Parameters	Values
Carrier Frequency (GHz)	9.65
System Bandwidth (MHz)	150
Incidence angle (°)	30
Slant range (Km)	1031.75
Observation number	24

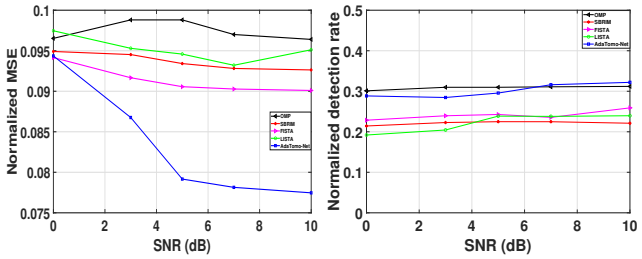
or overlaid double scatterers. We use OMP[2], SBRIM[9], FISTA[9], LISTA[3], and 10-layer AdaTomo-Net for TomoSAR imaging. To demonstrate the robustness to steering matrix perturbation of AdaTomo-Net, the steering matrix with perturbation is used to generate test samples in the same way mentioned in 4.4.1. The matrix without perturbations is used for TomoSAR imaging. Fig.3 shows that the performance of all algorithms degrades when the steering matrix is inaccurate, but the MSE of estimation of AdaTomo-Net is 2% lower than other algorithms when $\mathrm{SNR} > 5\mathrm{dB}$, which means that it has better performance in comparison with other algorithms.

4.1.3. Experiment with Simulated Facade

To demonstrate the AdaTomo-Net performance in multiple overlaid scatters case, we set a facade with continuous distributed scatters. The baseline configuration is the same as 4.4.1. The height of the facade is set to 30m. The SNR level is set to 3dB. We add quadratic PE the same way in 4.1.2. Fig.5 shows the SAR tomography results using SBRIM, LISTA, and 10-layer AdaTomo-Net. PGA preprocessing is performed before TomoSAR imaging with SBRIM. To evaluate the results more precisely, we use criteria $Accuracy$ and $Completeness$ which are introduced in [11]. Table 2 shows the criteria and thier values. According to [11], the lower $Accuracy$, the lower $Completeness$, the better the results. Thus we find that AdaTomo-Net has higher $Accuracy$ and $Completeness$ while keeping high computational efficiency at the same time.

4.2. Experiment with TerraSAR-X data

In this subsection, TerraSAR-X data are used to verify the performance of AdaTomo-Net on measured SAR data. PGA processing is performed before SBRIM. For AdaTomo-Net, we do not do PGA preprocessing. The imaging target is Barcelona Airport. Fig.5 shows the TomoSAR imaging results of SBRIM, LISTA and AdaTomo-Net. The time consumption of SBRIM, LISTA and AdaTomo-Net are 4559.87s, 1.79s, 1.86s respectively. We find that the TomoSAR imaging results of AdaTomo-Net have better completeness and continuity, which demonstrates the effectiveness of this network application on real SAR data.



(a) (b)

Fig. 3. TomoSAR imaging performance with point targets: (a) shows the Mean square error (MSE) of the imaging results w.r.t the different levels of SNR, (b) shows the detection rate of different algorithms w.r.t the different levels of SNR.

Table 2. The results evaluation of SBRIM, LISTA and AdaTomo-Net for SAR tomograhpy with simulated facade data.

	SBRIM	LISTA	AdaTomo-Net
Accuracy	22.1267	23.6294	**21.9736**
Completeness	25.3531	25.5998	**12.7453**
Times(s)	1637.47	**0.54**	0.65

5. CONCLUSION

In this paper, a deep learning approach AdaTomo-Net for TomoSAR imaging is proposed. We unrolled the sparsity-driven autofocus algorithm to a deep-unfolding network. Simulation experiment results show that the AdaTomo-Net can achieve better performance than classic CS algorithms when the steering matrix contains PE. Simulation experiments also verify that AdaTomo-Net is more effective in overlaid scatters separation than LISTA and classic CS algorithms under the condition of non-uniformed baseline configuration. In addition, the experiment with TerraSAR-X data demonstrates that AdaTomo-Net has competitive imaging performance compared with CS algorithm with PGA process, but is much more efficient. Therefore, we can see the application potential of AdaTomo-Net in future SAR tomography missions.

6. REFERENCES

[1] Xiao Xiang Zhu and Richard Bamler, "Superresolving sar tomography for multidimensional imaging of urban areas: Compressive sensing-based tomosar inversion,"

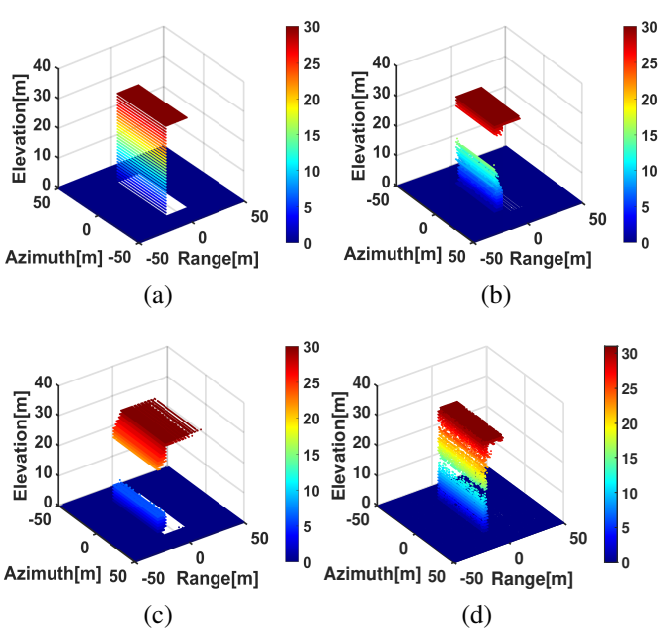


Fig. 4. TomoSAR imaging results with facade target: (a) is ground truth, (b) is SBRIM imaging result, (c) is LISTA imaging result, (d) is AdaTomo-Net imaging result

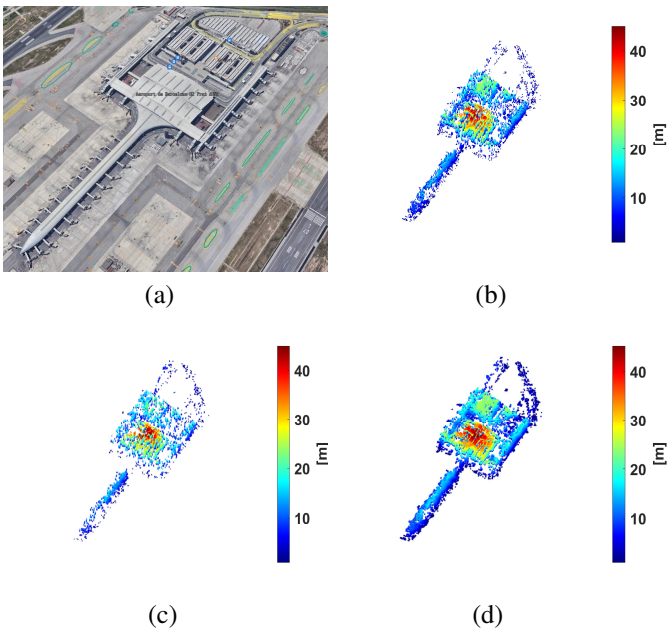


Fig. 5. TomoSAR imaging results with TerraSAR-X data: (a) is imaging scene given by Google Earth, (b) is SBRIM imaging result, (c) is LISTA imaging result, (d) is AdaTomo-Net imaging result

IEEE Signal Processing Magazine, vol. 31, no. 4, pp. 51–58, 2014.

[2] Jingkun Gao, Yu Ye, Shizhong Li, Yuliang Qin, Xunzhang Gao, and Xiang Li, "Fast super-resolution 3d sar imaging using an unfolded deep network," in *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*. IEEE, 2019, pp. 1–5.

[3] Kun Qian, Yuanyuan Wang, Yilei Shi, and Xiao Xiang Zhu, "Super-resolving sar tomography using deep learning," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 4810–4813.

[4] Aviad Aberdam, Alona Golts, and Michael Elad, "Ada-lista: Learned solvers adaptive to varying models," *arXiv preprint arXiv:2001.08456*, 2020.

[5] Hongliang Lu, Heng Zhang, Huaitao Fan, Dacheng Liu, Jili Wang, Xiangxing Wan, Lei Zhao, Yunkai Deng, Fengjun Zhao, and Robert Wang, "Forest height retrieval using p-band airborne multi-baseline sar data: A novel phase compensation method," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 175, pp. 99–118, 2021.

[6] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *Siam J Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[7] Shunjun Wei, Jiadian Liang, Mou Wang, Jun Shi, and Jinhe Ran, "Af-ampnet: A deep learning approach for sparse aperture isar imaging and autofocusing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–14, 2021.

[8] A. Aberdam, A. Golts, and M. Elad, "Ada-lista: Learned solvers adaptive to varying models," 2020.

[9] M. Wang, S. Wei, J. Shi, Y. Wu, and B. Tian, "Csr-net: A novel complex-valued network for fast and precise 3-d microwave sparse reconstruction," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. PP, no. 99, pp. 1–1, 2020.

[10] Fiona Muirhead, Bernard Mulgrew, Iain H Woodhouse, and D Greig, "Sparsity-driven autofocus for multipass sar tomography," in *SAR Image Analysis, Modeling, and Techniques XV*. International Society for Optics and Photonics, 2015, vol. 9642, p. 96420G.

[11] Clément Rambour, Loïc Denis, Florence Tupin, and Hélène M Oriot, "Introducing spatial regularization in sar tomography reconstruction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8600–8617, 2019.