SUPER-RESOLVING SAR TOMOGRAPHY USING DEEP LEARNING

Kun Qian[1], Yuanyuan Wang[2], Yilei Shi[1] Xiao Xiang Zhu[1,2]

(1) Data Science in Earth Observation, Technical University of Munich, Munich, Germany
(2) Remote Sensing Technology Institute, Germany Aerospace Center, Wessling, Germany

ABSTRACT

Synthetic aperture radar tomography (TomoSAR) has been widely employed in 3-D urban mapping. However, state-of-the-art super-resolving TomoSAR algorithms are computationally expensive, because conventional numerical solvers need to solve the l_2-l_1 mix norm minimization. This paper proposes a computationally efficient super-resolving TomoSAR inversion algorithm based on deep learning. We studied the potential of deep learning to mimic a conventional l_2-l_1 mix norm solver, i.e. iterative shrinkage thresholding algorithm (ISTA), and proposed several improvements of the complex-valued learned ISTA for TomoSAR inversion. Investigation on the super-resolution ability and estimator efficiency of the proposed algorithm shows that the proposed algorithm approaches the Cramer Rao lower bound (CRLB) with a computational efficiency more than 100 times better than the conventional solver.

Index Terms— SAR tomography, Super-resolution, Complex-valued neural network, Compressive sensing, deep learning

1. INTRODUCTION

Synthetic aperture radar tomography (TomoSAR) has been extensively applied in 3-D reconstruction in dense urban areas. The state-of-the-art TomoSAR inversion algorithms [1] [2] [3] are mostly based on compressive sensing (CS) techniques [4] since CS-based sparse reconstruction algorithms usually have strong super-resolution power and conduct to achieve the best performance using high-resolution SAR data like TerraSAR-X in the urban areas. However, due to the l_2-l_1 mixed norm minimization, CS-based algorithms are always computationally expensive for practical large-scale processing. Previous works investigated how to reduce the computational cost of CS-based algorithms, for instance, by integrating persistent scatterer interferometry (PSI) and TomoSAR [5] and by improving the efficiency of the optimizer [3]. However, the integration of PSI and TomoSAR [5] just reduces the percentage of pixels that need to be solved using CS-based methods by preclassifying the pixels. It did not fundamentally speed up the TomoSAR inversion procedure. In addition, although [3] boosts the solver of the l_2-l_1 mixed norm minimization, it is still hard to be extended to large-scale processing. In [6], the author investigated the potential of neural networks to detect single scatterer by formulating the inversion as a simple classification problem. Results showed that the algorithm [6] is able to achieve reasonable estimation accuracy in locating single scatterer. However, it cannot be applied to separate overlaid scatterers, which represents a general SAR tomographic inversion problem. In addition, there is no study so far about deep learning for TomoSAR in the super-resolution regime, which is a more challenging case.

In this paper, we proposed a novel deep learning based algorithm to address overlaid double scatterers detection and estimation in the super-resolving regime. We unroll iterative shrinkage thresholding algorithm (ISTA) [7] as a complex-valued feedforward neural network with side connection. The neural network is trained using datasets simulated according to specific spatial baselines. Once the neural network is well trained, it can be directly applied to further inference for SAR data with the same spatial baselines.

2. BACKGROUND AND PROBLEM FORMULATION

In the presence of noise $\boldsymbol{\epsilon}$, the discrete TomoSAR imaing model can be expressed as follows:

$$\mathbf{g} = \mathbf{R}\gamma + \boldsymbol{\epsilon}, \tag{1}$$

where $\mathbf{g} \in \mathbb{C}^{N\times 1}$ is the complex-valued SAR measurement vector with N elements g_n, $\mathbf{R} \in \mathbb{C}^{N\times L}$ is the steering matrix and $\boldsymbol{\gamma} \in \mathbb{C}^{L\times 1}$ denotes the discrete reflectivity profile vector. TomoSAR inversion is aimed at retrieving the reflectivity profile $\boldsymbol{\gamma}$ from the stacked SAR measurements for each range-azimuth cell, thus determining the number of overlaid scatterers and estimating their elevation. For TomoSAR problem with a sparse prior (e.g. in urban areas), the ideal solution of (1) is via l_0-norm minimization, which is, however, NP-hard. It is investigated in [1] that the reflected signal along the elevation direction is sparse enough so that the l_0-norm minimization can be approximated by l_1-norm minimization, which is able to give us nearly optimal estimates. Based on this fact, the reflectivity profile can be estimated as follows

$$\hat{\boldsymbol{\gamma}} = \arg\min_{\boldsymbol{\gamma}} \left\{ \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda\|\boldsymbol{\gamma}\|_1 \right\} \tag{2}$$

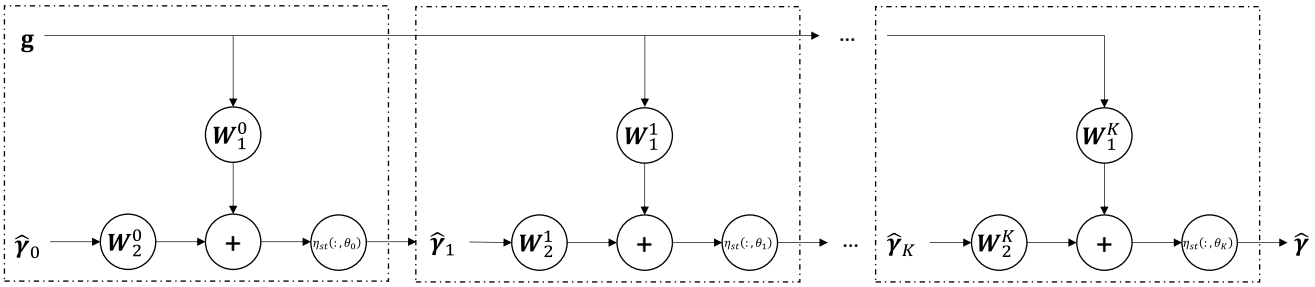


Fig. 1: Unfolded LISTA architecture. A K-layer LISTA unrolls the RNN and truncates it into K iterations, thus leading to a side-connected feedforward neural network.

λ is the factor to balance the data-fit and sparsity, it should be selected according to the noise level. Details about how to choose a proper λ can be found in [8].

3. SUPER-RESOLVING TOMOSAR VIA DEEP LEARNING

3.1. LISTA

ISTA [7] is a popular approach to solve equation (2), due to its simplicity and efficiency. Each iteration of ISTA is defined by

$$\hat{\boldsymbol{\gamma}}_i = \eta_{st}(\hat{\boldsymbol{\gamma}}_{i-i} + \beta \mathbf{R}^H \mathbf{b}_{i-1}, \theta_i)\,, \quad \text{with} \quad \mathbf{b}_i = \mathbf{g} - \mathbf{R}\hat{\boldsymbol{\gamma}}_i \tag{3}$$

where $\hat{\boldsymbol{\gamma}}_0 = \mathbf{0}$, β is a stepsize, η_{st} is the complex-valued version of the soft-thresholding function applied to each element of $\hat{\boldsymbol{\gamma}}_i$ and θ is a parameter controlling the soft-thresholding function to prune the elements with small magnitude. The complex-valued soft-thresholding function η_{st} is defined by

$$\eta_{st}(\hat{\boldsymbol{\gamma}}_i, \theta_i) = e^{i\cdot \arg(\hat{\boldsymbol{\gamma}}_i)} \max(|\hat{\boldsymbol{\gamma}}_i| - \theta_i, 0) \tag{4}$$

By rewriting (3) as the following form

$$\hat{\boldsymbol{\gamma}}_i = \eta_{st}\left\{\mathbf{W}_1^i \mathbf{g} + \mathbf{W}_2^i \hat{\boldsymbol{\gamma}}_{i-1}, \theta_i\right\} \tag{5}$$

where $\mathbf{W}_1^i = \beta \mathbf{R}^H$ and $\mathbf{W}_2^i = \mathbf{I} - \beta \mathbf{R}^{\mathrm{H}} \mathbf{R}$, we can find (5) is the basic form of the i^{th} layer in a recurrent neural network (RNN), if we view η_{st} as an activation function. Inspired by this fact, the learned ISTA (LISTA) is proposed in [9]. LISTA unrolls the RNN and truncates it into K iterations, thus leading to a K-layer neural network. Fig 1 illustrates the learning architecture of LISTA.

3.2. CV-LISTA for TomoSAR

To apply LISTA to solve TomoSAR inversion, we extend LISTA to complex-valued domain. Each neuron in complex-valued LISTA (CV-LISTA) is assumed to have two channels, which refer to the real and imaginary part of complex number, respectively. We execute the following adaptions to equation (5)

$$\tilde{\boldsymbol{\gamma}}_i = \eta_{st}\left\{\widetilde{\mathbf{W}}_1^i \tilde{\mathbf{g}} + \widetilde{\mathbf{W}}_2^i \tilde{\boldsymbol{\gamma}}_{i-1}, \theta_i\right\} \tag{6}$$

where

$$\widetilde{\mathbf{W}} = \begin{bmatrix} \Re(\mathbf{W}) & -\Im(\mathbf{W}) \\ \Im(\mathbf{W}) & \Re(\mathbf{W}) \end{bmatrix}, \tilde{\mathbf{g}} = \begin{bmatrix} \Re(\mathbf{g}) \\ \Im(\mathbf{g}) \end{bmatrix}, \tilde{\boldsymbol{\gamma}} = \begin{bmatrix} \Re(\hat{\boldsymbol{\gamma}}) \\ \Im(\hat{\boldsymbol{\gamma}}) \end{bmatrix}$$

and $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary operators, respectively. Accordingly, the matrix multiplication in the feedforward phase conforms to the complex number multiplication rule.

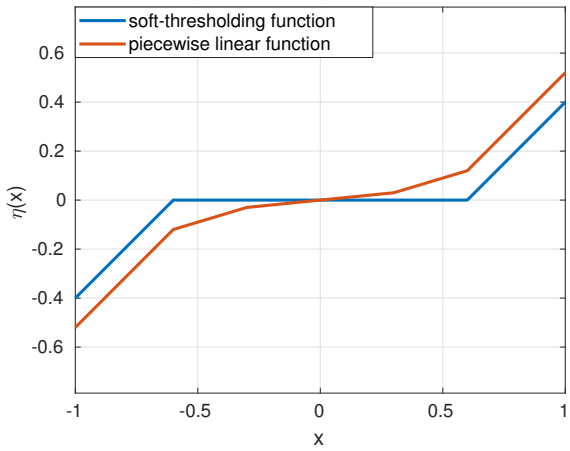


Fig. 2: Comparison between the piecewise linear function and soft-thresholding function

Due to the fact that the output of the intermediate layer in the CV-LISTA is generated exclusively on the output of the previous layer, minor errors are likely to be propagated and amplified in the upcoming layers. Furthermore, once the useful information is discarded, it is no longer possible to utilize the discarded information. Therefore, we give up the soft-thresholding function since pruning of small magnitude is very likely to result in information loss and replace it by

the piecewise linear function. The piecewise linear function is formally defined as follows:

$$\eta_{pwl}(\hat{\gamma}, \boldsymbol{\theta}) = \begin{cases} \theta_3 \hat{\gamma}, & |\hat{\gamma}| \leq \theta_1 \\ e^{i \cdot \arg(\hat{\gamma})}[\theta_4(|\hat{\gamma}| - \theta_1) + \\ \theta_3 \theta_1, & \theta_1 < |\hat{\gamma}| \leq \theta_2 \\ e^{i \cdot \arg(\hat{\gamma})}[\theta_5(|\hat{\gamma}| - \theta_2) + \\ \theta_4 (\theta_2 - \theta_1) + \theta_3 \theta_1], & |\hat{\gamma}| > \theta_2 \end{cases} . \tag{7}$$

Fig 2 compares the soft-thresholding function and the piecewise linear function we adopt. As we can see, the piecewise linear function just further shrinks the small magnitude rather than driving them to zero. It is conductive to maintenance of useful information. In our experiments, we find the piecewise linear function contributes to faster convergence and increasing the estimation accuracy.

4. EXPERIMENTS

4.1. Simulation setup

We verified the performance of the proposed algorithm on simulated data. We generated training and testing samples with a similar baseline setting in [2], i.e. 25 regularly distributed spatial baselines in the range -135m to 135m, which results in a Rayleigh resolution of about 42m. In the experiments, it is assumed that each simulated sample contains overlaid double scatterers with identical amplitude and phase, i.e., the worst case.

In the simulation, we generated samples with the normalized distance α between the double scatterers increasing from 0.1 till 1.2 with 12 samples and samples of different levels of SNR varying from 0 dB to 10 dB. The normalized distance α is defined as a ratio of the distance between the double scatterers to the Rayleigh resolution. For each pair of α and SNR, 0.2 million Monte Carlo trials were simulated. The testing data was generated in the same way, but at only four different SNR levels with $SNR = (0, 3, 6, 10)$ dB.

4.2. Performance assessment

The performance evaluation of a TomoSAR algorithm is twofold, i.e. the detection capability of multiple scatterers, and the efficiency and bias of the estimator. A good algorithm should have both high detection rate and high estimation accuracy. To better assess the performance, we introduced the "effective detection rate", which is a combination of the detection rate and the estimator accuracy. We define an effective detection in our experiments as follows: (1), the algorithm successfully detects two scatterers; (2), the estimates of their elevation should be within not only ± 3 CRLB but also $\pm 0.5 d_s$. d_s is the distance between the true positions of the double scatterers.

Fig 3 demonstrates the detection rate of the proposed algorithm w.r.t the normalized distance at different levels of SNR. It shows that the proposed algorithm achieves promising super-resolution power. To be specific, the proposed algorithm reaches about 80% effective detection rate at 0.5 Rayleigh resolution and 6 dB SNR, whereas the state-of-the-art method SL1MMER [2] algorithm reaches 85% under the same setting. For very noisy cases, i.e., $SNR = 0$ dB, the proposed method is still capable of separate overlaid double scatterers in super-resolving cases with high probability. For example, the proposed algorithm achieves about 60% of detection rate when the double scatterers are spaced by 0.8 Rayleigh resolution.

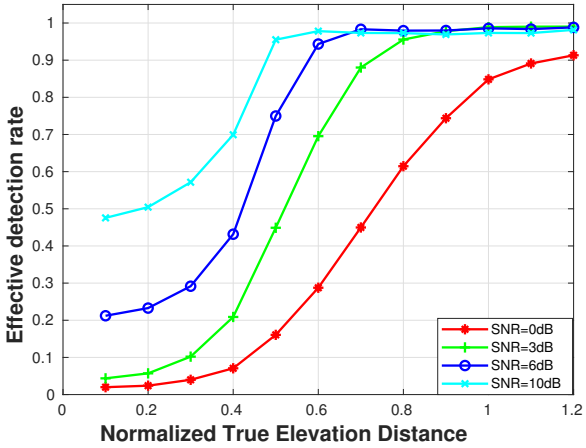
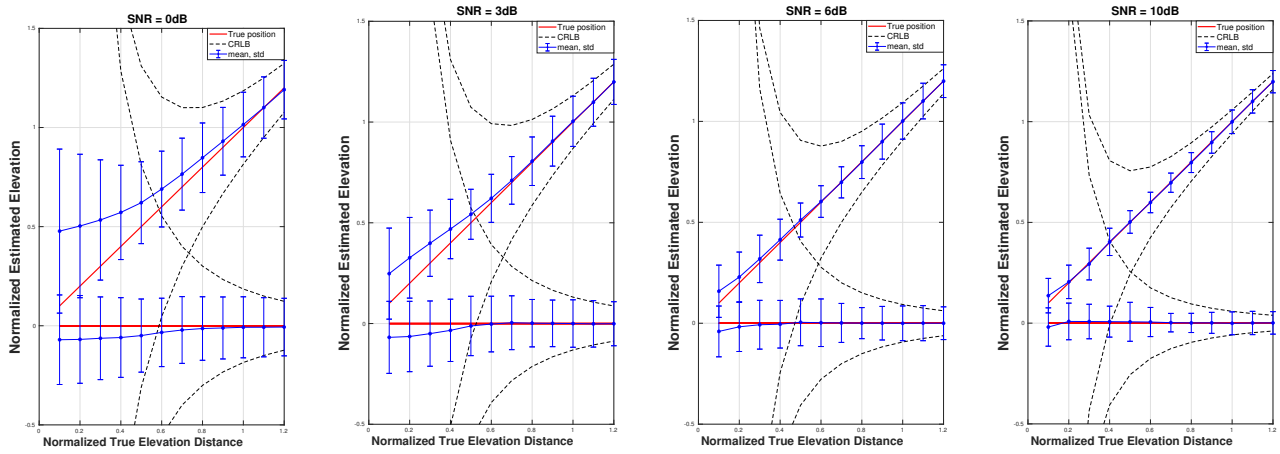


Fig. 3: Detection rate ρ_d as a function of the normalized elevation distance between the simulated double scatterers with SNR = (0,3,6,10) dB and N = 25 under 0.2 million Monte Carlo trials.

Fig 4 demonstrates the results of the facade-ground separation test at different SNRs, which is a well-known benchmark test for TomoSAR [1], [10]. In the figures, the elevation estimate of the simulated facade and ground are plotted w.r.t their normalized true elevation distance. Each blue dot has the mean of the estimates and the error bar depicting the corresponding standard deviation. The two red lines mark the true elevation of the simulated facade and ground. The dashed curves denote the true elevation $\pm 1\times$ CRLB. The difference between the dots and the corresponding red lines in y-axis indicates the bias of the estimates. As it is shown in the figures, the proposed algorithm provides high estimation accuracy with the standard deviation of elevation estimates within the CRLB in almost all cases, despite the fact that the estimates suffer from large bias when the double scatterers are extremely closely spaced. The bias of the estimates is approaching zero with the increasing normalized true elevation distance.



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

Fig. 4: Estimated elevation of simulated facade and ground, (a) $SNR = 0$ dB, (b) $SNR = 0$ dB, (c) $SNR = 6$ dB, (d) $SNR = 6$ dB. Each dot has the sample mean of all estimates as its y value and the correspond standard deviation as error bar.

5. CONCLUSION

In this paper, a deep learning based algorithm CV-LISTA for super-resolving TomoSAR inversion is proposed. We replace the conventional soft-thresholding function by the piecewise linear function to improve the performance of CV-LISTA for TomoSAR inversion in the super-resolution regime. Realistic simulations show that deep learning is capable of solving mixed L1- and L2-norm minimization in TomoSAR with accuracy approaching the CRLB. The proposed algorithm is also able to deliver competitive performance to the state of the art under the super-resolving scenario. Therefore, we see a high potential of applying deep learning for similar inverse problems in various applications.

6. REFERENCES

[1] Xiaoxiang Zhu and R. Bamler, "Tomographic sar inversion by l1-norm regularization – the compressive sensing approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, 2010.

[2] X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic sar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 247–258, 2012.

[3] Y. Shi, X. X. Zhu, W. Yin, and R. Bamler, "A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic sar," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 6148–6158, 2018.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] Yuanyuan Wang, Xiaoxiang Zhu, and Richard Bamler, "An efficient tomographic inversion approach for urban mapping using meter resolution sar image stacks," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1250–1254, 2014.

[6] A. Budillon, A. C. Johnsy, G. Schirinzi, and S. Vitale, "Sar tomography based on deep learning," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 3625–3628.

[7] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, 2004.

[8] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.

[9] Karol Gregor and Yann LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, Madison, WI, USA, 2010, ICML'10, p. 399406, Omnipress.

[10] Xiaoxiang Zhu and R. Bamler, "Very high resolution spaceborne sar tomography in urban environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 12, pp. 4296–4308, 2010.