URSI GASS 2021, Rome, Italy, 28 August - 4 September 2021

Deep learning algorithm for SAR autofocus

Wei Pu[(1)]
(1) University College London.

Abstract

The sparsity-driven technique is a widely used tool to solve the SAR imaging problem. However, it always encounters the sensitivity to motion errors. To solve this problem, this paper proposes a new deep neural network architecture, i.e., Sparse Auto-Encoder Network (SAE-Net). The proposed SAE-Net is designed to implement SAR imaging and autofocus simultaneously. In SAE-Net, the encoder transforms the SAR echo into imaging result, and the decoder regenerates the SAR echo using the obtained imaging result. The encoder is designed by the unfolded Alternating Direction Method of Multipliers (ADMM), while the decoder is formulated into the linear model. The experiments show that the proposed architecture outperforms other state-of-the-art autofocus methods in sparsity-driven SAR imaging applications.

1 Introduction

Over the past decade, synthetic aperture radar (SAR) has attracted increasing attention because of its wide applications in many civil and military fields. SAR image formation is a typical inverse problem[1, 2], wherein the reflectivity of the imaging scene can be recovered by exploiting its sparsity or compressibility through linear measurements[3]. However, such linear measurements are always subject to uncertainties originating from motion errors of the moving radar platform. Generally, the image quality will be significantly impaired by the motion errors in terms of artifacts, resolution losses and defocusing. Hence, many autofocus techniques have been proposed to eliminate the effect of motion errors.

Recently, several algorithms have been proposed to solve the autofocus problem in sparsity-driven SAR imaging methods. In [4], the effects of motion errors are modelled into azimuth phase errors, which are then estimated and compensated through a combination with sparsity-driven imaging. In reference [5], phase errors are divided into 3 categories (1-dimension phase error, 2-dimension separable phase error, and 2-dimension nonseparable phase error), which are then corrected by means of a nonquadratic regularization method. By contrast, reference [6] focuses mainly on the range-variance of the phase errors, which can compensate for the different azimuth phase errors corresponding to different range gates. Despite the existence of azimuth phase errors, residual range cell migration that stems from motion errors is taken into consideration in [7] as well.

These algorithms are capable of improving image quality. Nonetheless, motion errors haven't been fully formulated and corrected due to some approximations in the models. For instance, azimuth-variance of the azimuth phase errors and 2-dimensional spatial-variance of the residual range cell migration were neglected in the existing sparsity-driven SAR autofocus algorithms. This would degrade focusing quality especially in the case with large motion errors.

Under the above inspiration, we propose a Sparse Auto-Encoder Network (SAE-Net) in this paper, expecting to solve the SAR imaging and autofocus problem using an auto-encoder architecture. In the SAE-Net, encoder imitates the imaging operator. Decoder is formulated into a linear model. It is the inversion of the imaging operator, which is defined as a mapping from the imaging result to the SAR echo. In the proposed network, no more training data is needed, and the parameters in the SAE-Net are trained in a totally unsupervised manner.

We summarize the contributions of the paper as:

- **Encoder based on unfolded ADMM:** We design the encoder of the SAE-Net by reformulating the traditional Alternating Direction Method of Multipliers (ADMM) algorithm into a deep neural network. The network architecture is interpretable and it naturally inherits the prior knowledge such as the sparse transformation, imaging model and motion information of the radar platform. The output of encoder is imaging result.
- **Decoder based on linear model:** We design the decoder using a linear formulation, which is similar to a dense layer with zero-bias. The decoder is used to map the imaging result into the SAR echo and regenerate the SAR echo with a complete auto-encoder structure.
- **Losses:** Two types of losses, i.e., reconstruction loss and entropy loss, are designed to guide the training of SAE-Net. Among them, reconstruction loss of the auto-encoder is used to learn the parameters of the auto-encoder, while entropy losses are to promote an optimum imaging result.

- **Application on SAR data:** The effectiveness of the proposed autofocus algorithm has been verified on SAR data. The results show that the proposed algorithm rivals state-of-the-art in terms of compensation accuracy and image quality.

2 Sparsity-driven SAR imaging

In sparsity-driven SAR imaging problem, SAR echo $\mathbf{y}$ can be formulated as

$$\mathbf{y} = \mathbf{W}\mathbf{x}. \tag{1}$$

where $\mathbf{W}$ denotes the measurement matrix and $\mathbf{x}$ is the scattering coefficient matrix of imaging scene. To recover imaging scene $\mathbf{x}$ in (1), SAR imaging problem is always formulated into the Least Absolute Shrinkage and Selection Operator (LASSO) problem:

$$\min_x \frac{1}{2}\|\mathbf{y}-\mathbf{W}\mathbf{x}\|_2^2 + \lambda\|\mathbf{F}\mathbf{x}\|_1 \tag{2}$$

where $\mathbf{F}$ denotes the sparse transform matrix and $\lambda > 0$ is a scalar regularization parameter. To solve the LASSO problem in (2), Alternating Direction Method of Multipliers (ADMM) is widely used in the SAR imaging application. In ADMM, the problem in 2 can be solved in an iterative manner by following the updating rules in each iteration:

$$\mathbf{x}^{k+1} = (\mathbf{W}^H\mathbf{W} + \rho\mathbf{F}^H\mathbf{F})^{-1}(\mathbf{W}^H\mathbf{y} + \rho\mathbf{F}^H(\mathbf{z}^k - \mathbf{u}^k)) \tag{3}$$

$$\mathbf{z}^{k+1} = \mathscr{S}_{\lambda/\rho}\left(\mathbf{F}\mathbf{x}^{k+1} + \mathbf{u}^k\right) \tag{4}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{F}\mathbf{x}^{k+1} - \mathbf{z}^{k+1} \tag{5}$$

where $k \in \{1,2,\cdots,K\}$ represents kth iteration and superscript H corresponds to the Hermitian transpose operator. $\mathscr{S}(\cdot)$ is soft thresholding operator, which is defined as

$$\mathscr{S}_\alpha(x) = \begin{cases} x-\alpha, & x > \alpha \\ 0, & x = \alpha \\ x+\alpha, & x < \alpha \end{cases}. \tag{6}$$

3 SAE-Net for SAR autofocus

If movement information of the radar platform is known accurately, the measurement matrix $\mathbf{W}$ can be constructed precisely. The sparsity-driven SAR imaging techniques are capable of reconstructing $\mathbf{x}$ from echo $\mathbf{y}$ by utilizing the compressibility of $\mathbf{x}$. However, ADMM as well as other sparsity-driven SAR imaging algorithms are sensitive to motion errors, whereby actual measurement matrix $\mathbf{W}$ deviates from the nominal one. In this section, we propose SAE-Net to recover $\mathbf{x}$ in the presence of motion errors.

3.1 Structure

Auto-encoder is a family of neural network which is used to learn efficient data coding in an unsupervised fashion. It has an internal (hidden) layer that describes a code used to represent the input. This layer is constituted by two main parts: an encoder that maps the input into the code, and a decoder that maps the code to a reconstruction of the original input.

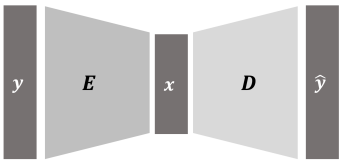

Figure 1. Structure of the CSAE-Net.

Fig. 1 depicts the main constituent blocks of our proposed SAE-Net based on the auto-encoder structure, where $\hat{\mathbf{y}}$ is the recovered imaging result of SAR echo, and $\mathbf{E}$ and $\mathbf{D}$ denote the encoder and decoder, respectively. In SAE-Net, the encoder $\mathbf{E}$ is a interpretable neural network, whereby prior information could be utilized, while the decoder $\mathbf{D}$ is formulated as linear model based on (1).

3.2 Encoder

To built a interpretable network, some model driven strategies have been proposed. Among them, the idea of unfolding[9] has recently drawn many attentions. Another advantage of unfolded network is that it is capable of exploiting the prior information when initialing the network. Inspired by this, the encoder in SAE-Net is designed by unfolding the ADMM introduced in the previous section. Similar to ADMM, the proposed unfolded ADMM network contains iteration blocks, where each iteration block is formed by $\mathbf{x}$, $\mathbf{z}$ and $\mathbf{u}$ layers.

In the ADMM algorithm, supposing that

$$\mathbf{H} = (\mathbf{W}^H\mathbf{W} + \rho\mathbf{F}^H\mathbf{F})^{-1}\mathbf{W}^H \tag{7}$$

and

$$\mathbf{B} = (\mathbf{W}^H\mathbf{W} + \rho\mathbf{F}^H\mathbf{F})^{-1}\rho, \tag{8}$$

(3) can be rewritten as

$$\mathbf{x}^{k+1} = \mathbf{H}\mathbf{y} + \mathbf{B}(\mathbf{z}^k - \mathbf{u}^k). \tag{9}$$

The idea of unfolding is that, instead of hand-coding $\mathbf{H}$ and $\mathbf{B}$, we use learnable parameters $\mathbf{H}^k$ and $\mathbf{B}^k$ to play the role of $\mathbf{H}$ and $\mathbf{B}$ at $(k+1)$th iteration. Therefore, the $\mathbf{x}$ update layer could be implemented by the computation $\mathscr{F}_x$ as

$$\mathbf{x}^{k+1} = \mathscr{F}_x(\mathbf{z}^k, \mathbf{u}^k, \mathbf{y}, \mathbf{H}^k, \mathbf{B}^k) = \mathbf{H}^k\mathbf{y} + \mathbf{B}^k(\mathbf{z}^k - \mathbf{u}^k). \tag{10}$$

Similarly, $\mathbf{z}$ update layer could be designed by the computation $\mathscr{F}_z$ as

$$\begin{aligned}\mathbf{z}^{k+1} &= \mathscr{F}_z(\mathbf{x}^{k+1}, \mathbf{u}^k, \mathbf{F}_1^k, \theta^k) \\ &= \text{sign}(\mathbf{F}_1^k\mathbf{x}^{k+1} + \mathbf{u}^k)\text{ReLU}\left(\left|\mathbf{F}_1^k\mathbf{x}^{k+1} + \mathbf{u}^k\right| - \theta^k\right),\end{aligned} \tag{11}$$

where $\theta^k = \lambda/\rho$ and $\mathbf{F}_1^k$ are set to be learnable parameters. $\mathrm{ReLU}(x) = \max(x,0)$ is one of the most commonly used nonlinear activation function in deep learning field, and

$$\mathrm{sign}(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases} \tag{12}$$

According to (5), the $\mathbf{u}$ update layer can be formulated as

$$\mathbf{u}^{k+1} = \mathscr{F}_u\left(\mathbf{x}^{k+1}, \mathbf{u}^k, \mathbf{F}_2^k\right) = \mathbf{u}^k + \mathbf{F}_2^k x^{k+1} - \mathbf{z}^{k+1}, \tag{13}$$

where $\mathbf{F}_2^k$ is learnable parameter as well.

Layers (10) to (13) can be formed into an iteration block, and one can form the encoder $\mathbf{E}$ by stacking each iteration block together. The unfolded network is equivalent to executing an ADMM iteration multiple times. Different from ADMM algorithm, the unfolded ADMM network doesn't need hundreds of iteration blocks. Here, we use 5 iteration blocks, and the corresponding structure is shown in Fig. 2.

3.3 Decoder

The decoder $\mathbf{D}$ is designed to reconstruct the SAR echo from the imaging result $\mathbf{x}$. According to (1), the corresponding reconstruction can be simply formulated into linear computation, which is implemented with a fully connected layer. Note that no activation function is employed and biases are set to be zero.

3.4 Initialization

In terms of encoder $\mathbf{E}$, the network parameters are initialized under the guidance of (7) and (8). Specifically, weights $\mathbf{H}^k$ and $\mathbf{B}^k$ associated with encoder $\mathbf{E}$ in (10) are initialized as the matrices $\mathbf{H}$ and $\mathbf{B}$ in (7) and (8), respectively. $\mathbf{F}_1^k$ and $\mathbf{F}_2^k$ are set to be the sparse transform matrix $\mathbf{F}$ in the initialization. In a similar manner, network parameters associated with $\mathbf{D}$ are initialized as matrix $\mathbf{W}$.

3.5 Learning strategy

Our strategy used to train the auto-encoder is developed and implemented based on various considerations:

- Firstly, we intend to minimize the reconstruction loss of the auto-encoder, which is defined as the difference between the reconstructed and original SAR echo. The reconstruction loss is given by

$$L_1 = \|\hat{\mathbf{y}} - \mathbf{y}\|_F. \tag{14}$$

- Secondly, we would like to optimize the imaging result $\mathbf{x}$ evaluated by some metric. Lots of image evaluation metrics, including entropy[8] and sharpness, have been applied to SAR imaging. Here, we choose entropy as it is more robust in assessing the quality of SAR images. It is widely accepted that a better focusing quality corresponds to a smaller value of entropy. Entropy is defined as

$$L_2 = E(\mathbf{x}) = \ln S - \frac{1}{S}\sum_{i=1}^{PQ} |\mathbf{x}_i|^2 \ln |\mathbf{x}_i|^2, \tag{15}$$

where S is the total energy of the imaging result $\mathbf{x}$,

$$S = \sum_{i=1}^{PQ} |\mathbf{x}_i|^2. \tag{16}$$

Therefore, we eventually learn the encoder and decoder by using the composite loss function:

$$L_{\text{total}} = L_1 + \lambda L_2, \tag{17}$$

where λ is the regularization parameter corresponding to the losses L_2. In this paper, we set $\lambda = 0.03$.

Finally, the learning process is completed through the Back Propagation for complex valued neural networks[10] at the learning rate of 0.000001. In particular, the encoder and decoder are learned simultaneously during the process, aimed to minimize the total loss L_{total}. Notably, the weights of the auto-encoder are learned in a totally unsupervised fashion while not any other training data is required.

4 Results

The effectiveness of the proposed algorithm is verified by the synthetic SAR data corresponding to extended target in the presence of motion errors. In the experiment, carrier frequency of the transmitted signal is 10GHz, bandwidth is 400MHz, and time width is 1.2μs. Synthetic aperture time is chosen to be 1.2s, radar platform velocity is $200m/s$, and pulse repetition frequency is 800Hz.

Fig. 3 depicts the imaging results corresponding to different autofocus algorithms. In detail, those obtained by ADMM without autofocus, by sparsity-driven autofocus in [6], sparsity-driven autofocus approach in [7] and the proposed algorithm are presented in Fig. 3 (a), (b), (c) and (d) respectively. In Fig. 3, azimuth is along the horizontal direction and range is along the vertical direction. The size of the imaging area is 1km×1km (Range×Azimuth). In Fig. 3 (a), the imaging result is seriously blurred, while the sparsity-driven autofocus algorithms in [6] and [7] promote a significant improvements in image quality. However, the imaging results in Fig. 3 (b) and (c) still suffer from severe image blurring. For instance, the rectangular house, which is located in the middle of the imaging scene, cannot be easily distinguished in Fig. 3 (b) and (c). Comparatively, the image quality of Fig. 3 (d) is improved by the proposed algorithm.

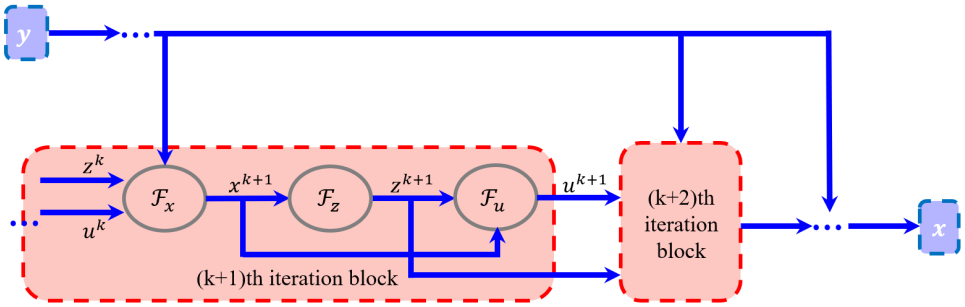

Figure 2. Structure of the Encoder **E**.

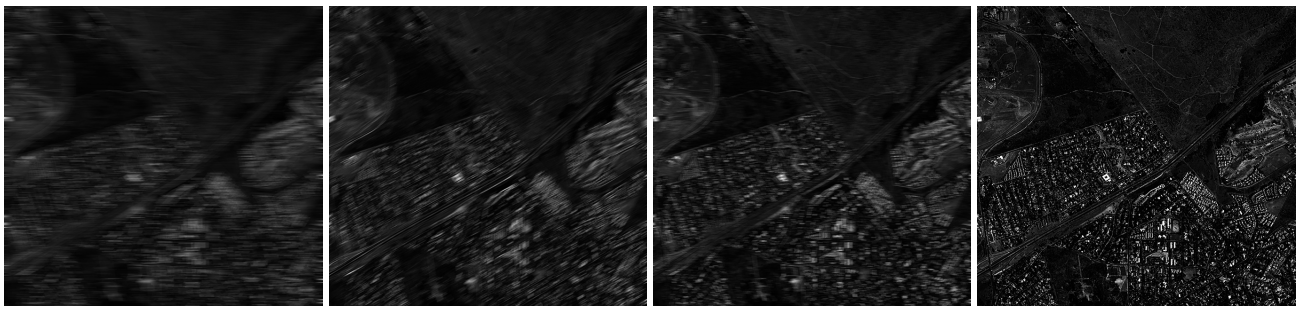

Figure 3. Columns 1 and 4 correspond to the imaging results by the ADMM algorithm without autofocus, sparsity-driven autofocus algorithm in [6] and sparsity-driven autofocus algorithm in [7], and the proposed algorithm, respectively.

5 Conclusions

For autofocus problem in SAR imaging, this paper proposes an auto-encoder structure named SAE-Net. In SAE-Net, the encoder is designed by unfolding the ADMM algorithm, which denotes the imaging processor. The decoder is a mapping from the imaging result to SAR echo. Both reconstruction losses and entropy losses are used to guide the training of the network. Experimental results verify the effectiveness of the proposed algorithm.

6 Acknowledgements

This work is supported by a Newton International Fellowship from the Royal Society.

References

[1] W. Pu and J. Wu, "OSRanP: A Novel Way for Radar Imaging Utilizing Joint Sparsity and Low-Rankness," *IEEE Trans. Comp. Imag.*, vol. 6, pp. 868–882, 2020.

[2] M. Cetin, I. Stojanovic, O. Onhon, K. Varshney, S. Samadi, W. C. Karl, and A. S. Willsky, "Sparsity-driven synthetic aperture radar imaging: Reconstruction, autofocusing, moving targets, and compressed sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 27–40, 2014.

[3] W. Pu, X.wang, J. Wu, Y. Huang and J. Yang, "Video SAR Imaging Based on Low-Rank Tensor Recovery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 188–202, 2021.

[4] S. Kelly, M. Yaghoobi, and M. Davies, "Sparsity-based autofocus for undersampled synthetic aperture radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 972–986, 2014

[5] N. O. Onhon and M. Cetin, "A sparsity-driven approach for joint SAR imaging and phase error correction," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2075–2088, 2012.

[6] Y. Chen, G. Li, Q. Zhang, Q. Zhang and X. Xia, "Motion Compensation for Airborne SAR via Parametric Sparse Representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 551–562, 2016.

[7] L. Yang, P. Li, S. Zhang, L. Zhao, S. Zhou and M. Xing, "Cooperative Multitask Learning for Sparsity-Driven SAR Imagery and Nonsystematic Error Autocalibration," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 5132–5147, 2020.

[8] W. Pu, "Joint Sparsity-based Imaging and Motion Error Estimation for BFSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1393 - 1408, March 2019.

[9] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ser. ICML'10. Madison, WI, USA: Omnipress, 2010, p. 399–406.

[10] H. Leung and S. Haykin, "The complex backpropagation algorithm," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 2101-2104, Sept. 1991