RFI Mitigation via Auto-Encoder Framework in SAR Data

Jiang Liu[1], Yunxuan Wang[1], Yuan Mao[1], Junli Chen[2], Yanyang Liu[2], Yan Huang[1]*

[1] State Key Lab of Millimeter Waves, Southeast University, China

[2] Shanghai Institute of Satellite Engineering, Shanghai, China

Introduction

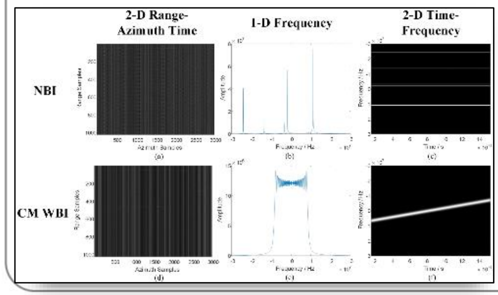
With the development of electronics and mobile wireless technology, the radio devices have an increasing demand for greater bandwidth. The radar system also requires larger bandwidth to obtain better range resolution, especially for imaging radar systems represented by synthetic aperture radar (SAR). Hence, radio frequency interference (RFI) becomes a major issue in accurate remote sensing by a SAR system, which greatly affect the SAR imaging and the following interpretation process.

During the past decades, lots of researchers have been working on how to effectively mitigate RFIs from collected SAR data. These methods can be approximately classified into three categories, i.e., non-parametric methods, parametric methods, and semi-parametric methods. Non-parametric methods mainly employ filters to filter out useful signal or subspaces to project the original data onto signal subspace. Parametric methods commonly estimate the parameters of RFIs, including narrowband and wideband models. Semi-parametric methods use the machine learning model, such as sparse recovery and low-rank recovery methods, to mitigate RFIs from measured data.

Recently, deep learning tools have excellent performance for nearly all the tasks in computer vision and natural language processing areas. For the supervised learning frameworks, they can extract the hierarchical features of targets in the images and further be employed to classify different targets. Previous methods, such as the deep convolutional neural network (DCNN), were used to mitigate both narrowband and wideband interferences in SAR images (W. Fan, F. Zhou, et. al., Remote Sens., 11(14), 1654, 2019). For the unsupervised learning frameworks, such as auto-encoder (Y. Wang, H. Yao, et. al., Neurocomputing, 184, 232-242, 2016) and generative adversarial network (GAN) (I. Goodfellow, J. Pouget-Abadi, M. Mirza, et al, Advances in neural information processing systems, 27, 2014), they can discover naturally intrinsic property of the data without any pre-assigned labels and are widely leveraged in the image denoising area. Also, the PCA-based unsupervised learning methods (M. Tao, J. Su, et. al, Remote Sens., 11, 2438, 2019) are widely used to mitigate RFIs in the corrupted SAR data. Therefore, we come out an idea that the unsupervised learning frameworks have great potential on separate the RFI and the useful data. In this paper, we use an auto-encoder framework, combining real and imaginary parts of complex radar data as two branches, to mitigate strong RFIs from corrupted SAR data. It brings another perspective for RFI mitigation via deep unsupervised learning approach.

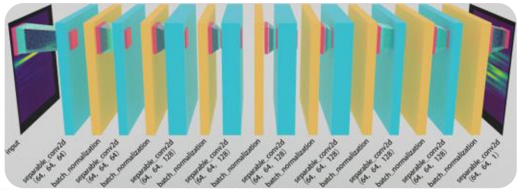
Signal Model

In SAR systems, the received signal of complex value SAR at range time t and azimuth time τ can be modeled as $x(t,\tau) = s(t,\tau) + I(t,\tau) + n(t,\tau)$, where $s(t,\tau)$, $I(t,\tau)$ and $n(t,\tau)$ denote the target received signal, RFI and additive noise, respectively. These RFIs can be classified into two categories according to the bandwidth: the narrow-band Interference (NBI) and the wide-band Interference (WBI). NBI can be expressed as $I_{NB}(t,\tau) = \sum_{k=1}^{K} a_k(t,\tau)\, exp\{j(2\pi f_k t + \phi_k)\}$, where $a_k(t,\tau)$, f_k and ϕ_k denote the complex amplitude, frequency and initial phase of the kth interference, respectively. Basically, the WBI have two forms, i.e., the chirp-modulated (CM) WBI and the sinusoidal-modulated (SM) WBI. Regards to the RFI, the CM WBI is typical and can be formulated as $I_{CM}(t,\tau) = \sum_{k=1}^{K} a_k(t,\tau)\, exp\{j(2\pi f_k t + \pi\gamma_k t^2)\}$, where $a_k(t,\tau)$, f_k and γ_k denote the complex amplitude, frequency and modulation frequency of the kth interference, respectively. Next, we introduce the short-time Fourier transform (STFT) to transform the signal vector into time-frequency domain. The basic idea of STFT is to perform Fourier transform (FT) to the part of the signal vector $y(t)$ via a smoothing window $h(t)$. Then the STFT can be formulated as $\mathrm{STFT}_y(t,f) = \int_{-\infty}^{\infty} y(t)h(t-\tau)e^{-j2\pi f\tau}d\tau$, where (t,f) indicates the coefficient in the time-frequency domain. Then the characteristics of both NBI and CM WBI in 2-D time domain, 1-D frequency domain, and 2-D time-frequency domain are shown in the left figure. As can be seen, the NBI performs like several strips in 2-D time-frequency domain, and the CM WBI looks like a slant line, showing linear frequency modulation property. But both RFIs perform irregularly in 2-D time domain and it is hard to recognize anything useful. This inspires us to mitigate the RFIs in 2-D time-frequency domain.



Proposed AutoEncoder Method

As analyzed before, the RFIs have distinct artifacts in range time-frequency domain. Therefore, it may be easily to remove RFIs in this domain rather than the image domain. The DCNN is able to extract features, textures, and other useful information in SAR images. For the RFI mitigation problem, we propose an RFI mitigation network (RMN) to mitigate strong RFIs, exploiting the advantages of deep learning methods. Previous works usually leverage supervised learning network, but the proposed RMN is based on the Auto-Encoder, one of the most typical unsupervised-learning networks. The input of the RMN is the time-frequency spectrum of the RFI-polluted signal and the other end of the RMN is that of the RFI-free signal. And the encoder extracts useful features and the decoder is used to recover the RFI-free signal. The RMN uses nine depth-wise separable convolution (DSC) layers (F. Chollet, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1800-1807, 2017) and eight batch normalization layers. The size of the convolutional kernel is 3 by 3, and the stride is 1. Next, we briefly introduce the DSC layer. The DSC layer can be regarded as a separation of a traditional convolutional layer. The traditional convolution layer is to take convolution of the whole data with a kernel, which has the same depth as the input. Suppose that the input size of the data is $M \times M \times N$ and the kernel size is $K \times K \times N \times P$, the parameter number is simply calculated as $K \times K \times N \times P$ without considering the bias number. The DSC layer actually has two convolutions. It first takes the convolution of each data channel with each kernel channel, and then $1 \times 1 \times N \times P$ kernel is used to finish another convolution. The total parameter number is $(K \times K \times N + N \times P)$, which is less than the original one. Therefore, with the same number of parameters, the network with DSC layers may go deeper than the one with traditional convolutional layers. The loss function is defined by the mean square error (MSE) as follows $L_{MSE}^{RMN} = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left[I_{ori}(m,n) - G_{RMN}\left(I_{input}(m,n)\right)\right]^2$, where the input size of the time-frequency spectrum is $M \times N$. Herein, I_{ori} and I_{input} are the RFI-free and RFI-polluted time-frequency spectra, respectively, and G_{RMN} denotes the network output. In this paper, we cut the whole time-frequency spectrum into multiple pieces with their size being 64 by 64. The time-frequency spectrum is complex data and we normalize the whole data and separate it into the real-part and imagery-part channels. Then the input data size is $64 \times 64 \times 2$ and the whole network is shown in the left figure. The batch is set to 20 and the learning rate is 0.001. We use the RMSProp as the optimization method and the initialization uses the Glorot method.



Experimental Results

In this section, we use the measured SAR data with simulated RFIs to provide enough data and support the RMN training. In order to quantitatively evaluate the performance of the proposed RMN, two metrics, interference suppression ratio (ISR) (M. Tao, F. Zhou, Z. Zhang, IEEE Trans. Geosci. Remote Sens., 54(1), 74-87, 2016) and signal distortion ratio (SDR) (F. Zhou, M. Tao, IEEE J. Sel. Top) are given as $ISR = 10\log_{10}\left(\frac{\sum|x|^2}{\sum|\hat{x}|^2}\right)$ and $SDR = 10\log_{10}\left(\frac{\sum|x_0-\hat{x}|^2}{\sum|x_0|^2}\right)$, where x_0, x and $\hat{x}$ denote the RFI-free, RFI-polluted and RFI-removal results, respectively. The larger ISR and the smaller SDR indicate better RFI mitigation performance. Herein, we consider two scenarios, one is the narrowband RFI and the other one is the wideband RFI. The input signal-to-interference-and-noise ratio (SINR) is -20 dB. We illustrate the time-frequency spectra of the narrowband RFI-polluted data and the RFI-free data, respectively, in Fig. (a) and (b). As can be seen, the interference is so strong that the RFI-free signal is totally submerged. Then we apply the proposed RMN to remove the narrowband RFI, and the result is shown in Fig. (c). The recovered result is quite similar with the RFI-free one. The ISR is 5.33 dB and the SDR is -12.34 dB, which indicates the good performance of interference mitigation.

Next, we test the proposed RMN on the CM wideband RFI case. The RFI-polluted and RFI-free time-frequency spectra are shown in Fig. (d) and (e). Due to the strong power of RFI, the RFI also dominates the whole spectrum with only a few useful signals can be seen. The recovered result is shown in Fig. (f) and it looks quite good. For this case, the ISR is 5.50 dB and the SDR is -12.79 dB. The proposed RMN has acceptable performance for strong RFI mitigation, including narrowband and wideband RFIs.