Radio Frequency Tomographic Reconstruction Based on Convolutional Neural Networks

Jia Li[1], Robert L. Ewing[2], and Xiaoping Shen[3]

[1]Dept. of Electrical and Computer Engineering, Oakland University, Rochester, MI 48309
[2]Sensors Directorate, Air Force Research Lab, WPAFB, OH 45433
[3]Department of Mathematics, Ohio University, Athens, OH 45701

***Abstract*—Convolutional neural network based radio frequency tomographic reconstruction is explored in this study. Due to the limited amount of sensors used in RF tomographic imaging system, analytical reconstruction such as filtered back-projection usually generates strong artifacts in the reconstructed images. The state of art algebraic reconstruction methods use prior knowledge to regularize reconstruction and reduce artifacts, but suffer from high computational complexity. Our study shows reconstruction based on convolutional neural network, a learning based approach, is effective in removing artifacts caused by limited number of sensors, and has low computational cost, which makes it suitable for real-time applications.**

***Index Terms*—radio frequency tomography, tomographic reconstruction, convolutional neural network**

I. Introduction

Radio frequency (RF) tomographic imaging is to estimate the dielectric properties of the field illuminated by RF pulses [1]. The RF transmitters and receivers are usually distributed to form a multi-static setting so that the received echoes carry the location and dielectric information of targets. Due to the spatial, frequency and waveform diversities, the waveform propagation channels are usually inhomogeneous. So the RF tomographic reconstruction can also be regarded as an information fusion process, where information collected by multiple RF sensors are fused to yield the best estimate of the dielectric property of the illuminated field.

The existing RF tomography reconstruction methods fall into two categories, analytical reconstructions and algebraic reconstructions. Filtered back-projection is the most commonly used analytical reconstruction algorithm. The advantage of analytical methods is the low computational complexity associated with them. However, due to limited number of sensors in a RF imaging system, only a limit number of echoes, which we also refer as "projections" in this paper, are recorded at the receivers. The back-projection of the echoes causes great amount of artifacts in the reconstruction result, which leads to very poor reconstruction quality. Algebraic reconstruction methods suffer less from artifacts phenomenon because they usually solve a system of equations that is built upon the data model from prior knowledge. The disadvantage of algebraic reconstructions is their high computational cost due to the iterative solution, which makes them impractical in real-time applications.

We propose a RF tomographic reconstruction method that is based on convolutional neural network (CNN). Studies of applying artificial neural networks in tomographic reconstruction problems have emerged in the last decade [2] [3]. The initial works have used Hopfield neural network as an optimization tool to minimize the difference between the measured projection data and the projections of reconstructed field. These neural network based methods share a common objective function with algebraic methods. The reconstruction time of these methods can be even larger than that of algebraic methods as they have to solve a nonlinear system instead of a linear system. We choose convolutional neural network (CNN) in our study for the low cost in training and execution. The objective function of network training calculates the difference between the inverse of the projection and the true image. The gradient of the objective function can be computed by chain rule. Therefore network parameters can be efficiently learned through gradient descent method. When network training is complete, the forward execution of a CNN only involves convolution and pointwise nonlinear operation, which makes the reconstruction fast and suitable for hardware acceleration. In our numerical experiment, the network is trained using synthesized data, where random ground truth images are generated, and projections are synthesized using known system model. In real practice, we can use known objects and their projections to train the network, or we can use an algebraic reconstruction method to obtain reconstructions and use them as ground truth to train the network.

As artifacts caused by limit number of sensors are determined by sensor locations and have fixed patterns, it's not easy to remove them by generic denoising approaches. The regularization used in algebraic reconstruction requires prior knowledge of the image, such as sparsity or object boundary, which is imposed into the objective function. The CNN based reconstruction can gain these knowledge naturally in the training phase instead of asking for it as an input.

In Section II, the system model of RF tomographic imaging is introduced. In Section III, we discuss the existing analytical and algebraic reconstruction methods, and propose the CNN based reconstruction method. After that, the results of numerical simulations are presented in Section IV. We conclude the paper and discuss future work in Section V.

978-1-5386-6557-2/18/$31.00 ©2018 IEEE

II. System Model

RF tomographic imaging system is usually composed of distributed RF sensors to achieve spatial and frequency diversities. The sensors can be homogeneous or inhomogeneous transmitters and receivers with multi-static setup. In the following derivation, we assume the pixels being imaged have isotropic reflection, waveforms are propagated in free space, and each sensor node is a duplex transceiver.

A. First order model

Let N be the number of pixels in the image. The grey-level of the j-th pixel, x_j, represents the pixel's dielectric reflectivity multiplied with the attenuation coefficient associated with the waveform propagation path. Considering a first order imaging system, the signal received at the m-th receiver is a linear combination of all the transmitted waveforms directly reflected by the field,

$$y_m(t) = \sum_{i=1}^{M} \sum_{j=1}^{N} x_j p(t - \tau_{ijm}) e^{-j2\pi f \tau_{ijm}} + n_m(t) \qquad (1)$$

where i is the index of the transmitter, $p(t)$ is the transmitted pulse, τ_{ijm} is the bistatic propagation delay of the path from the i-th transmitter to the j-th pixel, then to the m-th receiver, and $n_m(t)$ is an additive white Gaussian noise process. Please note the waveform propagated through direct path is not included in Eq. (1) since it is usually removed in the preprocessing step. By the first order model, we assume there is no secondary or higher order scatterers, or they are too weak compared to the first order reflectors. This assumption is applied to reduce the complexity of system model, and may not be valid in some real application. For example, when there are multiple targets in the field, a transmitted pulse can experience multiple level scattering before reaching the receiver. We discussed the formation of ghost targets and the model of multipath propagation in [4], where a sparse reconstruction method using dynamic dictionary was proposed for higher order systems.

Under the first order model, the set of transmitters and the m-th receiver uniquely establishes a linear projector which projects a sample $\mathbf{x}$ of the image space to generate a RF echo $y_m(t)$, i.e. a projection, in the measurement space. For a multi-static RF imaging system with M sensor nodes, a total of M projections can be obtained simultaneously.

B. Discrete model

The first order model described above can be discretized by sampling the transmitted pulse $p(t)$ and the received waveform $y(t)$. The image to be reconstructed can be stacked into a $N \times 1$ vector $\mathbf{x}$. This will lead to a discrete model for each projector $\mathbf{A}_m$,

$$\mathbf{y}_m = \mathbf{A}_m \mathbf{x} + \mathbf{n}_m \qquad (2)$$

where $\mathbf{y}_m$ is a column vector representing the sampled waveform received by the m-th receiver, and $\mathbf{A}_m = [\mathbf{p}_{1,m}, \ldots, \mathbf{p}_{N,m}]$ is a matrix whose column $\mathbf{p}_{j,m}$ is a sum of the delayed versions of the transmitted pulses reflected by the j-th pixel. $\mathbf{A}_m$ can be pre-calculated for each receiver in the imaging system. RF tomographic reconstruction is an inverse problem, which estimates the unknown vector $\mathbf{x}$ from the set of M observations $\{\mathbf{y}_m\}_{m=1}^{M}$.

III. Reconstruction

The traditional filtered back-projection (FBP) reconstruction is an analytical reconstruction method. It has low computational complexity as compared to iterative algebraic reconstruction method. However, image quality of FBP reconstruction is usually poor due to limited angle artifacts and noise in received waveforms. In RF tomographic imaging, the projector $\mathbf{A}$ is usually underdetermined due to limited number of sensors. Artifacts appear in FBP reconstruction of RF tomography as curved lines because back-projections are along ellipses with Tx and Rx as the focal points. To improve the quality of reconstructed images, various regularizations using prior knowledge of $\mathbf{x}$, such as sparsity or object boundary, have been proposed. The regularizer usually appear as a penalty term in the objective function:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \sum_{m=1}^{M} D(\mathbf{y}_m, \mathbf{A}_m \mathbf{x}) + \lambda \phi(\mathbf{x}), \qquad (3)$$

where D is a distance measure, λ is the non-negative weighting coefficient and ϕ is the penalty function. For example, the l_1 norm of $\mathbf{x}$ is commonly used to promote sparse solutions. The regularized optimization problem is usually solved through iterative methods, which have high computational cost and are not practical for real-time applications.

In recent years, artificial neural networks have been studied for solving inverse problems in imaging [5]. An artificial neural network can be used to model an unknown function $f : \mathbb{R}^n \to \mathbb{R}^m$. For a linear projector $\mathbf{A}$, given a set of images $\{\mathbf{x}_i\}_{i=1}^{n}$ and their corresponding projections $\{\mathbf{y}_i : \mathbf{y}_i = \mathbf{A}\mathbf{x}_i + \mathbf{n}\}_{i=1}^{n}$, the set of pairs $\{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{n}$ can be used to train a neural network and learn an inverse mapping $f \approx \mathbf{A}^{-1}$. Let $\boldsymbol{\theta}$ be the set of parameters of the neural network, network training will search the parameter space to find the optimal value of $\boldsymbol{\theta}$ that minimizes the distance between $\mathbf{x}_i$ and $f_{\boldsymbol{\theta}}(\mathbf{y}_i)$,

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} \sum_{i=1}^{n} D(\mathbf{x}_i, f_{\boldsymbol{\theta}}(\mathbf{y}_i)) + \phi(\boldsymbol{\theta}) \qquad (4)$$

where $\boldsymbol{\Theta}$ is the feasible parameter space, D is a distance measure, ϕ is a regularizer over the network parameters to prevent overfitting. It's worth noting that the direct result of network training is optimal network parameter $\boldsymbol{\theta}^*$ instead of reconstructed images. Once the network training is complete, $\boldsymbol{\theta}^*$ is applied to reconstruct an image from the measurements $\hat{\mathbf{x}} = f_{\boldsymbol{\theta}^*}(\mathbf{y})$.

In multi-static RF imaging, there are a set of linear projectors $\{\mathbf{A}_m\}_{m=1}^{M}$, where $\mathbf{A}_m$ is associated with the mth receiver. To reconstruct an image $\mathbf{x}$, the network shall take all the projections $\{\mathbf{y}_m : \mathbf{y}_m = \mathbf{A}_m \mathbf{x} + \mathbf{n}\}_{m=1}^{M}$ as input to minimize the difference between $\mathbf{x}$ and $f(\{\mathbf{y}_m\}_{m=1}^{M})$. So the network is no longer a model of the inverse mapping of a single

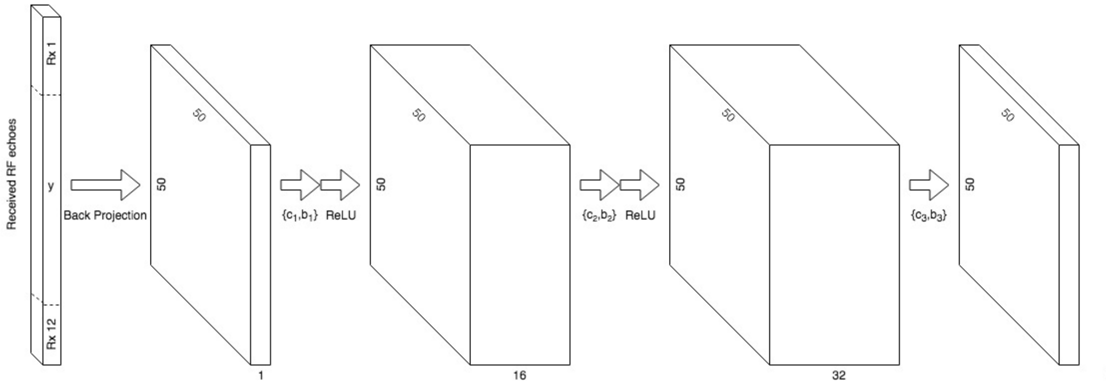


Fig. 1: Architecture of convolutional neural network for radio frequency tomographic reconstruction.

projector, but a fusion network which exploits information from all sensors in the learning process.

A. Network architecture

We have chosen convolutional neural network in our study for its success in other inverse problems in imaging and low computational cost [5]. In the design of network architecture, we have the option to map filtered back-projection as a layer in the network as suggested in [6], or take filtered back-projection results as the input to the network. Our investigation shows embedding FBP as a fully connected layer in the network has very high demand on memory and is not practical with the resource available to us. So the images reconstructed by filtered back-projection are used as the network input and the ground truth images are regarded as the labels of the data. With FBP as a preprocessing step, the network doesn't have to learn the physics of the imaging system, which saves the cost in network training. Figure 1 shows the architecture of the convolutional neural network in the proposed RF tomographic reconstruction. The RF echoes recorded at 12 receivers are back-projected to get the direct inverse, which is the initial input to the network. The network has 3 convolutional layers. Each layer is a set of linear filtering operations followed by a rectified linear unit, which is a nonlinear operation. The first layer contains 16 filters, the second layer contains 32 filters, and the last layer has 1 filter.

B. Network training

The convolutional neural network was trained using synthesized data. Our experiments consider reconstructions in two scenarios. The first scenario is there are multiple point targets in the field, while the second scenario is there are two large size targets in the field. For each scenario, a set of 1000 random images have been generated as the ground truth. For each random image, RF echoes at multiple receivers are synthesized using the first order system model. Euclidean distance is used as the distance measure in the objective function. The 1000 images and the associated RF echoes are evenly split to form the set of training data and the set of evaluation data. The stochastic gradient descent method is used to minimize the objective function. Overfitting of network is prevented by terminating the training when the performance on the evaluation set starts to decrease.

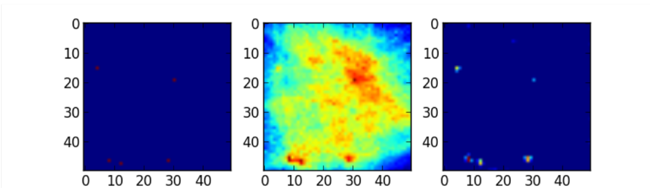
IV. Simulation Results

We present numerical simulations and reconstruction results obtained by the CNN based approach in this section. A two dimensional radio frequency imaging system, with the same setup as in [4], is simulated. A total of 12 RF sensor nodes are pseudo randomly distributed on a ring of 80m radius. Each node is a transceiver that can transmit and receive arbitrary waveforms in the frequency range of 200M-2G Hz. The dielectric property of the area inside of the ring is to be estimated through RF tomographic reconstruction. The image pixels are on a uniform 50×50 grid centered at the origin of the ring. The pixel resolution is 2m by 2m, which is chosen according to the bandwidth of the simulated pulse.

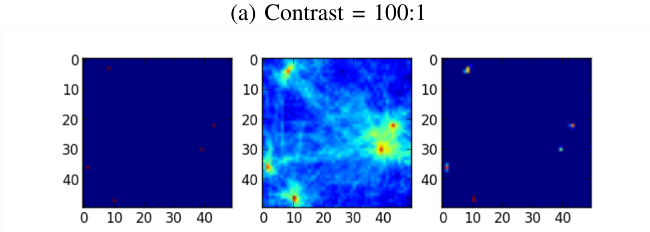
The simulated pulses are linear frequency modulated chirps with 40M Hz bandwidth at different central frequencies. A first order system model is used to synthesize the received waveforms. Two scenarios, including multiple point targets and two large targets, have been simulated using the described RF imaging system,

A. Multiple point targets

The first scenario has 5 point targets randomly located in the field. Different levels of image contrast have been simulated. Figure 2 shows the comparison of reconstruction results based on FBP and CNN in two different contrast levels. The top row is corresponding to a contrast of $100:1$, while the bottom row is corresponding to a higher contrast of $1600:1$. From left to right, the first column shows the ground truth image, the second column contains the FBP reconstruction results, and the third column shows the CNN reconstruction results. It can be seen that the quality of CNN based reconstructions is much better than the FBP reconstruction. Most of the artifacts contained in the FBP reconstruction have been removed except in the nearest neighborhood. While contrast level has great impact to the quality of FBP reconstruction, i.e. higher contrast level leads to better reconstruction, it has very little impact to the quality of CNN based reconstruction. We think this immunity to contrast level of CNN based reconstruction is due to its ability to gain knowledge of artifacts patterns associated with point targets, and generate customized filters to remove them.



(a) Contrast = 100:1



(b) Contrast = 1600:1

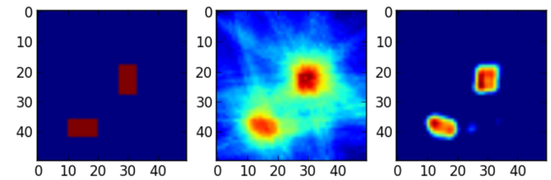
Fig. 2: Reconstruction of multiple point targets in different contrast.

B. Two large targets

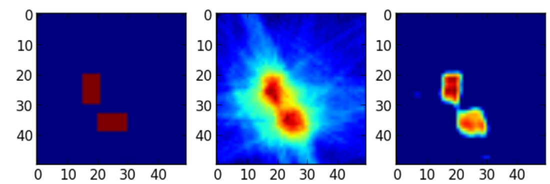
The second scenario has two large targets of rectangle shape randomly located in the field. Figure 3 shows the reconstruction results under the same conditions as those in the multiple point targets scenario. The first thing we observed is that contrast level doesn't have much impact to the quality of reconstruction in either FBP reconstruction or CNN based reconstruction. This is due to the negligibility of background reflectivity as compared to the strong reflectivity of large targets. Nevertheless, the CNN based reconstructions have removed most of the artifacts contained in the FBP reconstructions. The target boundaries in the CNN based reconstructions are not as sharp as those in the ground truth images, which is an expected result from the filtering function of CNN.

C. Limit number of sensors

We further investigated the performance of two reconstruction methods with varying number of RF sensors in the imaging system. The contrast of the ground truth image is kept as $1600:1$. Figure 4 shows how the signal to noise ratio (SNR) of the reconstructed images changes with the number of sensors. It's interesting to see that SNR of FBP reconstructions decreases when the number of sensors increases. This may be due to the higher amount of artifacts caused by increasing number of sensors. On the other hand, the SNR of CNN based reconstruction consistently increases with the number of sensors in both point targets and large targets scenarios. This demonstrates the capability of CNN to exploit the information gain brought by larger number of sensors.



(a) Contrast = 100:1



(b) Contrast = 1600:1

Fig. 3: Reconstruction of two large targets in different contrast.

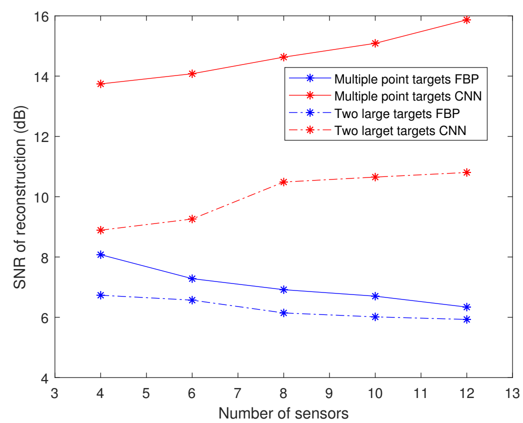


Fig. 4: RF tomographic reconstruction with different number of sensors.

V. Conclusions

We proposed a convolutional neural network based radio frequency tomographic reconstruction method in this paper. A 3-layer convolutional neural network is designed and trained for the task. After the network training is complete, applying the network to reconstruct RF tomography has a low computational cost. Numerical simulation shows the CNN based reconstruction can remove most of the artifacts contained in the filtered back-projection reconstruction. In the future study, we will explore the potential of convolutional neural network as a fusion network for data driven sensor fusion, for example, the fusion of radio frequency, electro-optical and infrared sensors in multi-static setup.

Acknowledgment

This research was supported by National Research Council Research Associateship Program and AFOSR grant FA9550-18-1-0287.

REFERENCES

[1] H. D. Griffiths and C. J. Baker, "Fundamentals of tomography and radar," in *Advances in Sensing with Security Applications Series 2*. Springer, 2006, pp. 171–187.

[2] R. Cierniak, "New neural network algorithm for image reconstruction from fan-beam projections," *Neurocomputing*, vol. 72, no. 13-15, pp. 3238–3244, 2009.

[3] D. M. Pelt and K. J. Batenburg, "Fast tomographic reconstruction from limited data using artificial neural networks," *IEEE Trans. on Image Processing*, vol. 22, no. 12, pp. 5238–5251, 2013.

[4] J. Li, R. L. Ewing, C. A. Berdanier, D. Wetzel, and X. Shen, "Sparse reconstruction of RF tomography with dynamic dictionary," in *2016 IEEE National Aerospace and Electronics Conference (NAECON) and Ohio Innovation Summit (OIS)*, 2016, pp. 391–395.

[5] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, Nov 2017.

[6] T. Würfl, F. Ghesu, V. Christlein, and A. Maier, "Deep learning computed tomography," in *Proc. of International Conference on Medical Image Computing and Computer Assisted Intervention*, 2016, pp. 432–440.