Plug-and-Play Reconstruction with 3D Deep Prior for Complex-Valued Near-Field MIMO Imaging

Okyanus Oral and Figen S. Oktem
Department of Electrical Engineering, Middle East Technical University (METU), Ankara, 06800 Turkey
Email: {ookyanus, figeno}@metu.edu.tr

***Abstract*—Near-field radar imaging systems are used in a wide range of applications, such as medical diagnosis, through-wall imaging, concealed weapon detection, and nondestructive evaluation. In this paper, we consider the inverse problem of reconstructing the three-dimensional (3D) complex-valued reflectivity distribution of the near-field scene from the sparse multiple-input multiple-output (MIMO) array measurements. Using the alternating direction method of multipliers (ADMM) framework, we formulate this problem by exploiting regularization on the magnitude of the complex-valued reflectivity distribution. We then provide a general expression for the proximal mapping associated with such regularization functionals operating on the magnitude of the complex-valued unknown. By utilizing this expression, we develop a computationally efficient plug-and-play reconstruction method that involves simple update steps both with analytical and deep priors. We illustrate the reconstruction performance of our approach with a 3D deep prior on a synthetic dataset. We also compare the result with the classical back-projection method and magnitude-total variation. Our results demonstrate that significant performance improvement can be achieved with learned 3D priors.**

***Index Terms*—Inverse problems, MIMO radar imaging, complex-valued optimization, plug-and-play, deep prior, compressed sensing.**

I. Introduction

Near-field radar imaging systems are used in a wide range of applications such as medical diagnosis, through-wall imaging, concealed weapon detection, and nondestructive evaluation [1]–[4]. Compared to the classical monostatic planar arrays (with colocated transmitter and receiver antennas), multiple-input multiple-output (MIMO) arrays offer reduced hardware complexity and cost for high-resolution imaging applications. As a result, there has been a growing interest in using MIMO arrays for near-field radar imaging [2], [4]–[6].

Reconstructing the three-dimensional (3D) complex-valued scene reflectivity from the sparse MIMO measurements is a highly ill-posed problem. Therefore the reconstruction quality greatly depends on the utilization of priors. For instance, the traditional direct inversion schemes, such as back-projection and range migration [3], [7], do not utilize prior information. Consequently, their performance substantially degrades with limited data. On the other hand, as motivated by the compressed sensing theory, the regularized reconstruction methods with sparsity priors such as Total-Variation (TV) and $\ell 1$ regularization offer promising imaging performance at compressive settings [8], [9]. However, sparsity-based reconstruction methods generally have iterative nature obtained using variable splitting methods such as the Alternating Direction Method of Multipliers (ADMM) [10]. As a result, these methods often have longer computation time than the direct inversion approaches, which is undesirable in real-time applications.

The regularized reconstruction methods in the near-field radar imaging literature enforce smoothness or sparsity directly on the complex-valued reflectivity distribution [8], [9], [11]. These methods are therefore built on the assumption that the scene reflectivity has locally correlated phase and magnitude. However, for many applications, the phase of the reflectivity at a particular point can be more accurately modeled as random and uncorrelated with the phase at other points [12], [13]. This is because phase shift can occur when imaging rough surfaces and also at the air/target interface due to the electrical properties of materials [12]. Accordingly, enforcing regularization only on the magnitude of the complex-valued scene reflectivity can improve imaging performance [13]–[15].

With the recent developments in deep learning, learned reconstruction methods emerged as powerful alternatives to the regularized reconstruction methods with hand-crafted analytical priors [16], [17]. These methods can utilize deep neural networks to learn data-driven priors. The state-of-the-art learned reconstruction methods incorporate the forward model of the imaging system into the reconstruction process to exploit physics-based information. Learned Plug-and-Play (PnP) regularization and unrolling-based approaches are examples of physics-based learned reconstruction methods [16], [18]. These approaches replace the proximal mappings on the variable splitting steps with deep neural architectures [17], [18]. Despite the recent success of these methods, most of the approaches are developed for 2D and real-valued image reconstruction problems [16]–[18]. Furthermore, there is no study on such methods for near-field MIMO imaging where we encounter a 3D complex-valued image reconstruction problem.

In this paper, we develop a novel and efficient PnP method for reconstructing the 3D complex-valued reflectivity distribution of the near-field scene from sparse MIMO measurements. Due to the random phase nature of the scene reflectivities in various applications, we formulate the image formation problem by exploiting regularization on the magnitude of the reflectivity function. To use in variable splitting frameworks such as ADMM, we provide a general expression for the proximal mapping associated with such regulariza-

This work is supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under grant 120E505.

tion functionals operating on the magnitude. By utilizing this expression, we develop a computationally efficient plug-and-play reconstruction method that results in simple update steps for each variable. Since our ADMM-based method also works with learned priors, we utilize a 3D deep denoiser that can jointly exploit range and cross-range correlations. The developed PnP approach provides a unified framework to efficiently handle arbitrary regularization on the magnitude of a complex-valued unknown, which appears to be missing in the previous related radar imaging works [14], [15]. We demonstrate the effectiveness of our approach in a compressive near-field imaging scenario using a synthetic dataset consisting of extended targets and compare the performance with analytical approaches.

II. Observation Model

Under Born approximation, we can express the discrete forward model that relates the near-field MIMO radar measurements to the discretized reflectivity distribution of the scene as follows [3], [11]:

$$y(x_T, y_T, x_R, y_R, k) = p(k) \sum_{x,y,z} \frac{e^{-jk(R_T+R_R)}}{4\pi R_T R_R} s(x,y,z) \quad (1)$$

Here $y(x_T, y_T, x_R, y_R, k)$ denotes the measurement obtained using the transmitter at $(x_T, y_T, 0)$ and the receiver at $(x_R, y_R, 0)$. The measurements are expressed in the temporal Fourier domain with f denoting the temporal frequency and $k = \frac{2\pi f}{c}$ denoting the frequency-wavenumber. $p(k)$ represents the Fourier transform of the transmitted pulse. The 3D complex-valued reflectivity distribution of the scene is denoted by $s(x,y,z)$ and, R_T and R_R denote the distances of the corresponding transmitter and receiver antennas to the voxel at (x,y,z), i.e. $R_T = \sqrt{(x_T-x)^2 + (y_T-y)^2 + z^2}$, $R_R = \sqrt{(x_R-x)^2 + (y_R-y)^2 + z^2}$.

The discrete model in (1) can be expressed in matrix-vector product form as follows:

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{n} \quad (2)$$

where $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the observation matrix for the ordered image vector $\mathbf{s} \in \mathbb{C}^N$ and the corresponding measurement vector $\mathbf{y} \in \mathbb{C}^M$. The vector $\mathbf{n} \in \mathbb{C}^M$ represents uncorrelated Gaussian noise with standard deviation σ_n.

The matrix $\mathbf{A}$ is a rectangular matrix with $N \gg M$ where M equals the product of the number of transmitter antennas, the number of receiver antennas, and the number of frequency steps, while N is the number of voxels in the discretized reflectivity image. The $(m,n)^{th}$ element of the observation matrix represents the contribution of the nth voxel to the mth measurement and can be expressed as follows:

$$\mathbf{A}_{m,n} = p(k_m) \frac{e^{-jk_m(R_{Tmn}+R_{Rmn})}}{4\pi R_{Tmn} R_{Rmn}}. \quad (3)$$

III. Plug-and-Play Reconstruction Method

A. Formulation of the Inverse Problem

In the inverse problem, the goal is to estimate the three-dimensional complex-valued reflectivity field, $\mathbf{s}$, from the radar measurements, $\mathbf{y}$. Due to the random phase nature of the scene reflectivities, we formulate the inverse problem using a regularization function, $\mathcal{R}(|\cdot|)$, that only operates on the reflectivity magnitudes:

$$\min_{\mathbf{s}} \mathcal{R}(|\mathbf{s}|) \text{ subject to } \|\mathbf{y} - \mathbf{A}\mathbf{s}\|_2 \leq \epsilon \quad (4)$$

Here, ϵ is a parameter that should be chosen based on the variance of the measurement noise (i.e. $\sqrt{M \cdot \sigma_n^2}$).

B. Variable Splitting and ADMM

To solve this regularized inverse problem, we first convert the constrained problem to an unconstrained one using a penalty function and then apply variable splitting as follows:

$$\begin{aligned} &\min_{\mathbf{s},\mathbf{v}_1,\mathbf{v}_2} \iota_{\|\mathbf{y}-\mathbf{v}_1\|_2 \leq \epsilon}(\mathbf{v}_1) + \mathcal{R}(|\mathbf{v}_2|) \\ &\text{subject to } \mathbf{A}\mathbf{s} - \mathbf{v}_1 = 0\ ,\ \mathbf{s} - \mathbf{v}_2 = 0 \end{aligned} \quad (5)$$

Here the indicator function $\iota_{\|\mathbf{y}-\mathbf{v}_1\|_2 \leq \epsilon}(\mathbf{v}_1)$ takes value 0 if the constraint in (4) is satisfied and $+\infty$ otherwise, whereas $\mathbf{v}_1$, $\mathbf{v}_2$ are the auxiliary variables.

We solve the optimization problem in (5) using C-SALSA approach [19] by obtaining the associated augmented Lagrangian form and then alternatively minimizing over $\mathbf{s}$, $\mathbf{v}_1$, and $\mathbf{v}_2$. In the corresponding ADMM framework, the minimization over $\mathbf{s}$ corresponds to solving a least-squares problem with the following normal equation:

$$(\mathbf{A}^H\mathbf{A} + \kappa\mathbf{I})\mathbf{s}^{l+1} = \mathbf{A}^H(\mathbf{v}_1^l + \mathbf{d}_1^l) + \kappa(\mathbf{v}_2^l + \mathbf{d}_2^l) \quad (6)$$

where the superscript l is the iteration number, $\kappa \in \mathbb{R}^+$ is a hyper-parameter that needs to be adjusted, and $\mathbf{d}_1$, $\mathbf{d}_2$ denote the dual variables for $\mathbf{A}\mathbf{s}$ and $\mathbf{s}$. Noting that solving this equation using matrix inversion is impractical due to the large size, we use the conjugate-gradient (CG) method to iteratively find the solution of the normal equation.

The minimization over $\mathbf{v}_1$ corresponds to projection of $\mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l$ onto ϵ-radius hyper-sphere with center $\mathbf{y}$, and is computed as:

$$\mathbf{v}_1^{l+1} = \mathbf{y} + \begin{cases} \epsilon \frac{\mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l - \mathbf{y}}{\|\mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l - \mathbf{y}\|_2}, & \text{if } \|\mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l - \mathbf{y}\|_2 > \epsilon \\ \mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l - \mathbf{y}, & \text{if } \|\mathbf{A}\mathbf{s}^{l+1} - \mathbf{d}_1^l - \mathbf{y}\|_2 \leq \epsilon \end{cases} \quad (7)$$

At last, the minimization over $\mathbf{v}_2$ corresponds to the proximal operator for the regularization function, $\mathcal{R}(|\cdot|)$, that operates on the magnitude of a complex-valued vector:

$$\mathbf{v}_2^{l+1} = \mathbf{\Psi}_{\alpha\mathcal{R}(|\cdot|)}(\mathbf{s}^{l+1} - \mathbf{d}_2^l) \quad (8)$$

where $\mathbf{\Psi}_{\alpha\mathcal{R}(|\cdot|)}(.)$ is the respective proximal operator given by

$$\mathbf{\Psi}_{\alpha\mathcal{R}(|\cdot|)}(\mathbf{p}) = \arg\min_{\mathbf{v}} \alpha\mathcal{R}(|\mathbf{v}|) + \frac{1}{2}\|\mathbf{v} - \mathbf{p}\|_2^2 \quad (9)$$

with $\alpha \in \mathbb{R}^+$ tuning the amount of regularization per iteration.

C. Handling Phase for Regularization on the Magnitude

In this section, we provide the solution of (9) by approaching this optimization problem as a joint minimization over the phase and magnitude of $\mathbf{v}$. The complex-valued vectors are re-written as a product of a complex-valued unitary diagonal phase matrix and a real-valued non-negative magnitude vector:

$$\mathbf{v} = \mathbf{\Phi_v}|\mathbf{v}|,\ \mathbf{p} = \mathbf{\Phi_p}|\mathbf{p}| \tag{10}$$
$$\text{where } \mathbf{\Phi_v} = \text{diag}(e^{j\angle(\mathbf{v})}),\ \mathbf{\Phi_p} = \text{diag}(e^{j\angle(\mathbf{p})})$$

With this change of variables, the problem in (9) becomes

$$\arg\min_{|\mathbf{v}|,\mathbf{\Phi_v}} \alpha\mathcal{R}(|\mathbf{v}|) + \frac{1}{2}\| \begin{bmatrix} |\mathbf{v}| \\ |\mathbf{p}| \end{bmatrix} \|_2^2 - |\mathbf{v}|^T \mathfrak{Re}\{\mathbf{\Phi_v^*\Phi_p}\}|\mathbf{p}| \tag{11}$$

For any given $|\mathbf{v}|$ and $|\mathbf{p}|$, the minimizing phase of (11) must maximize the elements of $\mathfrak{Re}\{\mathbf{\Phi_v^*\Phi_p}\}$. That is, the proximal mapping of a function that operates on the magnitude of a complex-valued vector must directly pass the phase values of the proximal point. After inserting this optimal phase solution to (11), optimization for the magnitude becomes:

$$\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}(|\mathbf{p}|) = \arg\min_{|\mathbf{v}|} \alpha\mathcal{R}(|\mathbf{v}|) + \frac{1}{2}\||\mathbf{v}| - |\mathbf{p}|\|_2^2 \tag{12}$$

where $\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}$ is now the Moreau proximal mapping associated with the regularization function $\mathcal{R}(\cdot)$, and corresponds to denoising of the magnitude of the proximal point $\mathbf{p}$ when the noise variance is α. A similar derivation for the scalar-valued case appears in [20].

This result suggests that the proximal update step in (9) can be computed as

$$\mathbf{\Psi}_{\alpha\mathcal{R}(|\cdot|)}(\mathbf{p}) = \mathbf{\Phi_p}\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}(|\mathbf{p}|) \tag{13}$$

which corresponds to denoising the magnitude of $\mathbf{p}$ using $\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}$ and merging the denoised magnitude with the unprocessed phase of $\mathbf{p}$. In our work, we choose $\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}$ as easy-to-compute proximal operators and compute the multiplication with the diagonal phase matrix using element-wise products.

Equation (13) decouples the magnitude and phase updates, and enables us to extend the real-valued proximal operators to the complex-valued optimization problem in (9). To summarize, each iteration of our PnP reconstruction algorithm consists of three computationally efficient update steps. The first step is the scene update given in (6) and carried out using the CG algorithm. The second step is the projection given in (7) and efficiently computed using scaling operations. The last step is the denoising step in (8) which is shown to be equivalent to (13), and corresponds to directly passing the phase but denoising the magnitude of the proximal point.

IV. 3D Deep Denoiser for PnP Regularization

Following the success of convolutional neural networks (CNN) on denoising [16], [17], [21], we train and deploy a deep CNN-based Gaussian denoiser for the third step of our PnP approach. Our denoiser is a 3D U-net developed based on [22] and is shown in Fig. 1. This denoiser replaces the proximal operator $\mathbf{\Psi}_{\alpha\mathcal{R}(\cdot)}$ in (13).

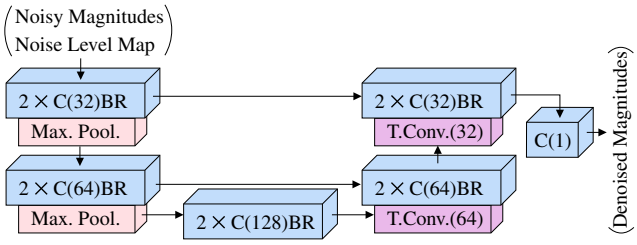


Fig. 1: Network architecture of the proposed deep denoiser.

The proposed denoiser is a 3-level encoder-decoder architecture with repeated 3D convolutional blocks (C) followed by batch normalization (B) and ReLU (R). On each level, max pooling (Max. Pool.) is used to reduce the spatial size of the input tensor by a factor of 2 in each dimension and transposed convolution blocks (T.Conv.) are used to increase by 2. At each decoding level, the output of the transposed convolution block is concatenated with the encoder outputs. The concatenated outputs are then fed to the respective decoding blocks. A single-channel 3D convolution block follows the last decoding block. The number of output channels of all convolutional blocks is indicated inside parentheses in the figure.

The input of the U-net is the 3D reflectivity magnitude that will be denoised and the 3D noise level map. The noise level map adjusts the amount of denoising in our non-blind denoiser network and its values are set to the constant α in (12). The output of the U-net is the 3D denoised reflectivity magnitude.

V. Experiments

A. Simulation Setting

To illustrate the performance of the developed approach, we consider an application in microwave imaging with a commonly used 2D Mill's Cross array [3]. Physical dimensions of the antenna array and the scene of interest are shown in Fig. 2. On the MIMO array, 13 transmitter and 12 receiver antennas uniformly span the diagonals of the 2D antenna plane in a cross configuration. The frequency, f, is swept between 4 GHz to 16 GHz with 15 uniform steps.

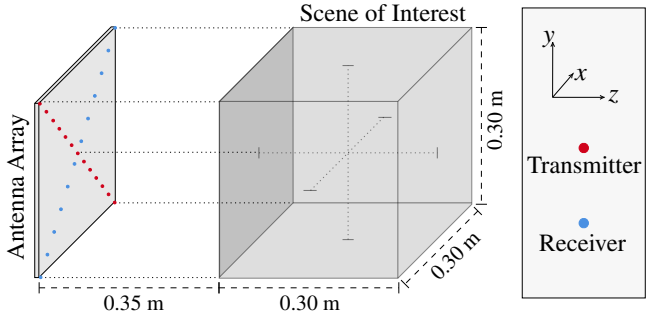


Fig. 2: Illustration of the imaging setting.

The theoretical resolution of this MIMO system is 2.5 cm in the cross-range directions, and 1.25 cm in the down-range direction [3]. We discretize the reflectivity distribution inside

the scene of interest using a sampling period of 1.25 cm for x, y directions, and 0.625 cm for z direction. For the scene of interest in Fig. 2, this results in an image cube of size $25 \times 25 \times 49$. Hence the compressive imaging task is to reconstruct this image cube with $\sim 8\%$ data (M/N).

Because a large experimental dataset is not available for microwave imaging, we use a synthetic dataset [23]. This dataset consists of randomly generated 800, 100, and 100 complex-valued image cubes of size $25 \times 25 \times 49$ for training, testing, and validation, respectively. The magnitude of these image cubes is obtained by randomly generating 15 points within the cube and passing them through a 3D Gaussian filter to obtain volumetric objects. The scene magnitudes are normalized via the sigmoid function to the range $[0, 1]$. Moreover, each scene voxel has random phase that is independently drawn from a uniform distribution between 0 and 2π.

B. Training of the 3D Deep Denoiser for PnP Regularization

We train our deep denoiser by minimizing the mean squared error between the ground truth magnitudes and Gaussian noise added magnitudes on 800 training scenes. We use a batch size of 16 with the maximum number of epochs set as 2000. We utilize Adam optimizer [24] with an initial learning rate of 10^{-3}, and drop the learning rate by a factor of 10 if the validation loss does not improve for 25 epochs. We stop the training if the validation loss does not improve for 50 epochs. At the end of the training, we use the network weights that corresponds to the minimum validation loss.

At each iteration of training, a new Gaussian noise realization is added to each ground truth magnitude by randomly and uniformly choosing the noise standard deviation, σ_ν, from the interval $[0, \frac{50}{255}]$. In addition, the constant noise level map is formed using the value of noise variance, $\alpha = \sigma_\nu^2$, and concatenated to the noisy magnitude. Hence the denoiser network learns to denoise the reflectivity magnitudes in a non-blind manner by providing the noise level map as input.

Training takes approximately 15 minutes on NVIDIA GeForce RTX 3080 Ti GPU using PyTorch 1.12.0 with CUDA Toolkit 11.6.0 in Python 3.10.6.

C. Results

We simulate the measurements using the model in (2) with a signal-to-noise ratio (SNR) of 30 dB, where SNR is defined as $10 \log_{10} \left(\frac{\|\mathbf{As}\|_2^2}{M \cdot \sigma_n^2} \right)$. We comparatively evaluate the performance with the back-projection algorithm, isotropic total variation (TV) and learned PnP regularization. For TV regularization on the reflectivity magnitude, we utilize Chambolle algorithm [25] for denoising the magnitude in (13). We call our magnitude TV and learned PnP-based C-SALSA algorithms TVCG-CSALSA and NNCG-CSALSA, respectively.

We initialize C-SALSA iterations using $\mathbf{s}^0 = \frac{\mathbf{A}^H \mathbf{y}}{\max(|\mathbf{A}^H \mathbf{y}|)}$. At each $\mathbf{s}$-update-step, the conjugate gradient algorithm is run for 5 inner iterations. For NNCG-CSALSA, we choose the number of C-SALSA iterations as 30. For TVCG-CSALSA, we run Chambolle algorithm for 5 inner iterations for each $\mathbf{v}_2$-update-step. We terminate TVCG-CSALSA when the relative change in magnitudes $\frac{\| |\mathbf{s}|^{l+1} - |\mathbf{s}|^l \|_2}{\| |\mathbf{s}|^l \|_2}$ drops below $5 \cdot 10^{-4}$.

For TVCG-CSALSA, the iterations converge to a solution for a sufficiently large κ in (6), which is set in our case to $\kappa = 5 \cdot 10^4$. For NNCG-CSALSA, the value of κ is optimized using the validation dataset, and is set to $\kappa = 5 \cdot 10^2$. For both of the algorithms, the α parameter in (13) determines the amount of regularization in each iteration. We also search for the optimal value of α using the validation dataset, and set it to $3 \cdot 10^{-3}$ and $7 \cdot 10^{-2}$ for TVCG-CSALSA and NNCG-CSALSA, respectively.

Although the reconstructed reflectivity distribution is complex-valued, its phase is random and does not contain any useful information. For this reason, we evaluate the performance of the methods on the normalized magnitudes. We use 3D peak signal-to-noise ratio (PSNR) and 3D structural similarity index measure (SSIM) [26] as quantitative performance metrics. The average performance of the algorithms on the test dataset consisting of 100 image cubes are given in Table I.

TABLE I: Average Performance on the Test Dataset

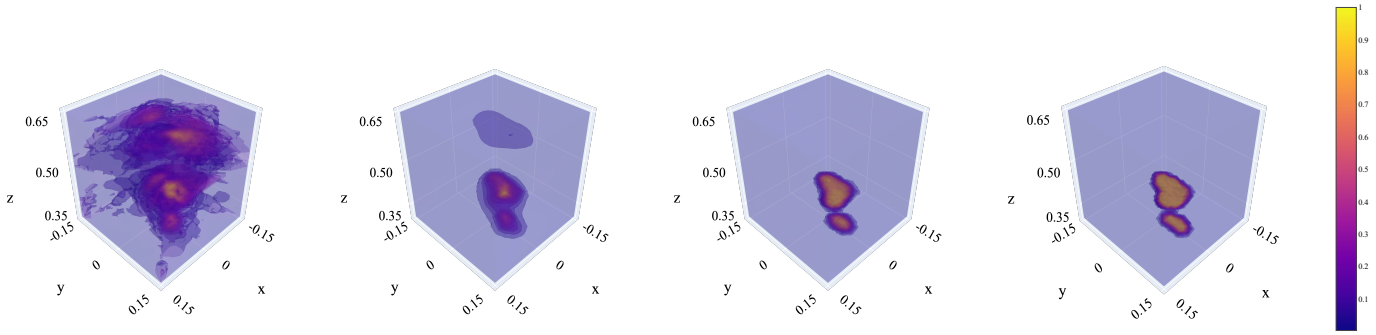
Algorithm	**Iterations**	**Seconds**	**PSNR** (dB)	**SSIM**
Back-Projection	1	0.26	21.71	0.235
TVCG-CSALSA	361	21	24.26	0.296
NNCG-CSALSA	30	2.3	29.30	0.851

In terms of average PSNR/SSIM, the developed PnP method NNCG-CSALSA with 3D deep prior significantly outperforms other methods by achieving the highest average performance of 29.30/0.851. TVCG-CSALSA has the second-highest performance with 24.26/0.296, whereas the performance of the back-projection is 21.71/0.235. Although the back-projection algorithm is the fastest with 0.26 seconds of average runtime, its reconstructions contain many artifacts. On the other hand, the developed PnP method with deep prior provides good image quality with a runtime of only 2.3 seconds. TVCG-CSALSA is the slowest method among the tested algorithms with an average runtime of 21 seconds.

To visually compare the results, we also provide in Fig. 3 the magnitude of the reconstructions for a sample test image. As seen, NNCG-CSALSA provides the best image quality with its reconstruction closely resembling the ground truth. While the back-projection result largely fails with many reconstruction artifacts, TVCG-CSALSA reconstruction contains over-smoothing where the two separate clusters on the ground truth now appear connected. Moreover, an artifact is also present on the upper part of the TVCG-CSALSA result. On the other hand, NNCG-CSALSA successfully reconstructs the ground truth image cube consisting of two separate clusters in this highly compressed observation setting with $\sim 8\%$ data.

VI. Conclusion

In this paper, we developed a novel PnP method for reconstructing the 3D complex-valued reflectivity distribution of a scene from sparse near-field MIMO measurements.



(a) Back-Projection Output (b) TVCG-CSALSA Output (c) NNCG-CSALSA Output (d) Ground Truth

Fig. 3: Example reconstructions from the test dataset. From left to right, reconstruction PSNRs (SSIMs) on the test image are 20.07 (0.147), 25.60 (0.308), and 30.86 (0.900).

The experimental results suggest that significant performance improvement can be achieved with learned 3D priors that jointly exploit correlations along range and cross-range dimensions. Our approach provides a generalizable means for effectively handling arbitrary regularization on the magnitude of a complex-valued unknown and can be applied to other radar image formation problems (including SAR) with similar prior information.

REFERENCES

[1] F. Fioranelli, S. Salous, and X. Raimundo, "Frequency-modulated interrupted continuous wave as wall removal technique in through-the-wall imaging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 10, pp. 6272–6283, 2014.

[2] S. S. Ahmed, A. Schiessl, F. Gumbmann, M. Tiebout, S. Methfessel, and L.-P. Schmidt, "Advanced microwave imaging," *IEEE Microwave Magazine*, vol. 13, no. 6, pp. 26–43, 2012.

[3] X. Zhuge and A. G. Yarovoy, "Three-dimensional near-field MIMO array imaging using range migration techniques," *IEEE Transactions on Image Processing*, vol. 21, no. 6, pp. 3026–3033, 2012.

[4] E. Anadol, I. Seker, S. Camlica, T. O. Topbas, S. Koc, L. Alatan, F. Oktem, and O. A. Civi, "UWB 3D near-field imaging with a sparse MIMO antenna array for concealed weapon detection," in *Radar Sensor Technology XXII*, vol. 10633. SPIE, 2018, pp. 458–472.

[5] X. Zhuge and G. Yarovoy, Alexander, "Study on two-dimensional sparse MIMO UWB arrays for high resolution near-field imaging," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 9, pp. 4173–4182, 2012.

[6] M. B. Kocamis and F. S. Oktem, "Optimal design of sparse MIMO arrays for near-field ultrawideband imaging," in *2017 25th European Signal Processing Conference (EUSIPCO)*. IEEE, 2017, pp. 1952–1956.

[7] Y. Álvarez, Y. Rodriguez-Vaqueiro, B. Gonzalez-Valdes, F. Las-Heras, and A. García-Pino, "Fourier-based imaging for subsampled multistatic arrays," *IEEE Transactions on Antennas and Propagation*, vol. 64, no. 6, pp. 2557–2562, 2016.

[8] E. A. Miran, F. S. Oktem, and S. Koc, "Sparse reconstruction for near-field MIMO radar imaging using fast multipole method," *IEEE Access*, vol. 9, pp. 151 578–151 589, 2021.

[9] S. Li, G. Zhao, H. Li, B. Ren, W. Hu, Y. Liu, W. Yu, and H. Sun, "Near-field radar imaging via compressive sensing," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 2, pp. 828–833, 2015.

[10] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[11] F. S. Oktem, "Sparsity-based three-dimensional image reconstruction for near-field MIMO radar imaging," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 5, pp. 3282–3295, 2019.

[12] D. Munson and J. Sanz, "Image reconstruction from frequency-offset Fourier data," *Proceedings of the IEEE*, vol. 72, no. 6, pp. 661–669, 1984.

[13] M. Çetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 623–631, 2001.

[14] M. B. Alver, A. Saleem, and M. Çetin, "Plug-and-Play synthetic aperture radar image formation using deep priors," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 43–57, 2021.

[15] H. E. Güven, A. Güngör, and M. Cetin, "An augmented Lagrangian method for complex-valued compressed SAR imaging," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 235–250, 2016.

[16] G. Ongie, A. Jalal, C. A. M. R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE Journal on Selected Areas in Information Theory*, 2020.

[17] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-Play image restoration with deep denoiser prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 10, pp. 6360–6376, 2021.

[18] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-Play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, 2013, pp. 945–948.

[19] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2010.

[20] J. Fessler, "EECS 598-006, optimization methods for signal and image processing and machine learning, chapter 5 proximal methods," vol. 20, 2021. [Online]. Available: https://web.eecs.umich.edu/~fessler/course/598/l/

[21] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3929–3938.

[22] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.

[23] I. Manisali, *Deep Learning-Based Reconstruction Methods for Near-Field MIMO radar imaging*, Ankara. [Online]. Available: https://open.metu.edu.tr/handle/11511/98756

[24] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[25] A. Chambolle and T. Pock, "A first-order primal-dual algorithm for convex problems with applications to imaging," *Journal of mathematical imaging and vision*, vol. 40, pp. 120–145, 2011.

[26] A. K. Moorthy and A. C. Bovik, "Efficient motion weighted spatio-temporal video SSIM index," in *Human Vision and Electronic Imaging XV*, B. E. Rogowitz and T. N. Pappas, Eds., vol. 7527, International Society for Optics and Photonics. SPIE, 2010, p. 75271I. [Online]. Available: https://doi.org/10.1117/12.844198