SOLVING 3D RADAR IMAGING INVERSE PROBLEMS WITH A MULTI-COGNITION TASK-ORIENTED FRAMEWORK

Xiaoling Zhang, Xu Zhan, Xiangdong Ma, Jun Shi, Shunjun Wei, and Tianjiao Zeng

University of Electronic Science and Technology of China

ABSTRACT

This work focuses on 3D Radar imaging inverse problems. Current methods obtain undifferentiated results and don’t meet the task’s specific demands well. For example, biased scattering energy may be acceptable for screen imaging but not for scattering diagnosis. To address this issue, we propose a new task-oriented imaging framework. The imaging principle is task-oriented through an analysis phase to obtain task’s demands. The imaging model is multi-cognition regularized to embed and fulfill demands. The imaging method is designed to be generalized, where couplings between cognitions are decoupled and solved individually with approximation and variable-splitting techniques. Tasks include scattering diagnosis, person screen imaging, and parcel screening imaging are given as examples. Experiments on data from two different 3D radar systems indicate that the proposed framework outperforms the current ones in task-dependent information retrieval.

Index Terms— 3D Imaging, radar imaging, task-oriented, multi-cognition, inverse problem

1. INTRODUCTION

3D radar imaging technology has shown potential values in multiple areas like urban structures’ health, military equipment scattering diagnosis, person screen imaging, and parcel screening imaging [1-4]. In practice, the forward process from the image \( \alpha_{x,y,z} \) to the demodulated echoes \( y_{x_p,y_p,r} \) can be expressed as

\[
y_{x_p,y_p,r} = \sum_{x,y,z} \alpha_{x,y,z} s_i(r - r_{x,y,z}) + n_{x_p,y_p,r} \quad (1)
\]

where \( x_p, y_p \) and \( r \) are sampling indexes for three directions, \( x, y, \) and \( z \) are image indexes. \( s_i(r - r_{x,y,z}) \) is the delayed transmitting signal. \( n_{x_p,y_p,r} \) is the noise-and-clutter. In practice, only limited bandwidth echoes can be measured, which undermines the backward/inverse process, and only partial information can be retrieved. In this work, we focus on retrieving target-dependent information. A new multi-cognition task-oriented framework is proposed, as illustrated in Fig.2. It’s designed with considerations through the whole process from three articulated aspects, including the imaging principle, model, and the imaging method.
2.1. Task-oriented imaging principle

The most applied framework is matched-filtering, treating the target and environment equally. To differentiate target from environment, the sparse-oriented approach is proposed. However, it only considers the difference between target and environment, which remains fixed regardless of the task. As radar systems develop to measure larger bandwidth, higher resolution images with more details become available. This presents an opportunity to meet differentiated task demands, as shown in Figure 2’s left column, known as the task-oriented principle. The distinct feature is an additional analysis phase compared to current approaches. The task is analyzed to obtain specific demands beyond target and environment aspects. These resulting demands vary in number and content based on different tasks. Three tasks are given as examples:

Scattering diagnosis: Locate main point-scatters on the target and determine their normality based on energy and position in the image. This task demands accurate retrieval of main scatters’ amplitudes and positions, along with suppressing clutter and noise to avoid false alarms.

Person screen imaging: Detect and recognize concealed dangerous objects on the human body, relying on the target’s shape. Retrieval of main point-scatters and relatively weak distributed-scatters (which form the shape) is necessary.

Parcel screen imaging: Similar to personal screening, this task aims to detect dangerous objects, but in a more complex environment. The target is inside a box or bag, creating strong interference. The third additional demand is to suppress interferences from the surroundings.

2.2. Multi-cognition regularized imaging model

For simplicity and generalization, (1) can be expressed in the matrix-vector form, as shown in the middle column in Fig.2. Commonly, a consistency regularization term exists that guarantees measured echoes and estimated scatters are inconsistent with the electromagnetic wave propagation rule. In the model of the sparse-oriented framework, the target is enhanced through an additional sparse regularization term, which is mathematically described by the target’s $l_1$ norm.

As multiple demands exist for a task, each corresponds to a cognition term. Intuitively, we can extend the sparse-oriented model into one with multi-cognition regularization terms.

\[
\mathbf{x} = \arg\min_{\mathbf{x}} \frac{1}{2} \| \mathbf{y} - \mathbf{A} \mathbf{x} \|_2^2 + \sum_i \beta_i g_i(\mathbf{x}) \quad (2)
\]

Where $\mathbf{x}$ and $\mathbf{y}$ are the folded target image and echo image, respectively. $g_i(\mathbf{x})$ is the $i$th cognition regularization term, and $\beta_i$ is the weight. These terms are mathematical descriptions to fulfill the different demands. The following sub-section will give examples corresponding to the former three tasks.

2.3. Generalized imaging method

The model in (2) seems to be reasonable at first glance. However, three critical insights should be pointed out. First, as the target and echo images have to be reshaped into 1D vectors to fit the model, their original 3D tensor structures are destroyed. This structure integrity loss may have little influence on the sparsity regularization because it does not describe the mutual information of pixels. However, to fulfill multi-cognition regularizations, this loss can’t be ignored. For example, shape cognition usually relies on the original 3D space. Second, to fit the folded $\mathbf{x}$, the observation matrix $\mathbf{A}$ would be relatively large, causing large storage and computation burden. Third, there are obvious couplings between regularization terms. As these regularization terms are variable for the task-oriented imaging framework, the coupling problem can hinder the generalization of the method.

To address the first and second issues, we introduce the approximation technique, on the basis of (2), we modify it as follows.

\[
\mathcal{X} = \arg\min_{\mathcal{X}} \frac{1}{2} \| \mathcal{Y} - f_{\text{eq}}(\mathcal{X}) \|_F^2 + \sum_i \beta_i g_i(\mathcal{X}) \quad (3)
\]

Where $\mathcal{X}$ and $\mathcal{Y}$ are original 3D tensor forms of target image and echo image, respectively. $f_{\text{eq}}(\cdot)$ is a generalized version of the observation matrix $\mathbf{A}$ that is functionally equivalent. We dub it the echo generation operator, which is a corresponding adjoint operator of the imaging operator $f_{\text{eq}}(\mathcal{Y})$. Specifically, they adhere to the relationship as follows.

\[
f_{\text{eq}}(\mathcal{X}) \approx \mathcal{Y}, \quad f_{\text{eq}}(\mathcal{Y}) = f_{\text{eq}}^\dagger(\mathcal{Y}) \approx \mathcal{X} \quad (4)
\]

Where $f_{\text{eq}}^\dagger$ denotes the adjoint operator. The imaging operator is a combination of steps originating from the matched-filtering imaging method. Taking the RMA as an example.

\[
f_{\text{eq}}(\mathcal{Y}) = \text{ifft}_3(\text{fft}_3(\mathcal{Y}) \odot P_{\zeta}) \quad (5)
\]
Where \( \text{fft}_3 / \text{iff}_3 \) denotes two-dimensional DFT/IDFT on each frontal slice \( Y(:, :, t) \) along the range direction, and \( P_C \) denotes the filter tensor that compensates the corresponding phase [6]. \( \odot \) denotes the point-wise Hadamard product. Then, the echo generation operator can be derived through their corresponding inverse steps as follows.

\[
\text{filt} (A) = \text{iff}_3 (\text{fft}_3 (A) \odot P_C) \tag{6}
\]

Where \( P_C^* \) is the conjugate of \( P_C \).

To address the third issue, we introduce the variable splitting technique, where alternative variables are utilized to replace the original ones in an alternating direction method of multipliers (ADMM) framework. Specifically, it can be formed as follows.

\[
\begin{align*}
& \min_{x, z, \gamma} \frac{1}{2} \| Y - \text{filt}(A) \|^2_F + \sum_i \left( \beta_i g_i(z_i) + \frac{\gamma_i}{2} \| x_i - z_i + \frac{1}{\gamma_i} D_i \|^2_F \right) \tag{7}
\end{align*}
\]

Where \( \gamma \) is the penalty parameter, and \( D_i \) is the multiplier. And (7) can be solved iteratively. Within the \( k \)th iteration block, the process contains steps as follows.

\[
\begin{align*}
& A^{k+1} = \arg \min_{A} \frac{1}{2} \| Y - \text{filt}(A) \|^2_F + \frac{1}{2} \| x - A \|^2_F \\
& D_i^{k+1} = \frac{1}{\gamma_i} D_i^k + \lambda \varepsilon_{i+1} - Z_i^{k+1} \\
& Z_i^{k+1} = Z_i^k + (1/\gamma_i) D_i^k + \frac{1}{\gamma_i} g_i(z_i)
\end{align*}
\]

Taking the formerly mentioned tasks as examples. 1) Scattering diagnosis. The first demand can be fulfilled by the target’s \( l_p \) norm regularization with the threshold-based function as the proximal operator. And the second one can utilize the noise-and-clutter’s \( l_2 \) norm regularization, with its operator being the shrinkage function. 2) Person screen imaging. Without the high precision requirement, the first demand can utilize a less complex \( l_1 \) norm with a shrinkage-threshold function proximal operator. And the second demand can be met through Shearlet-Transform which is capable of characterizing directional distributed-scatters. And the regularization term can be the \( l_1 \) norm in the transformed domain. 3) Parcel screen imaging. Apart from the above two, the kernel norm of interferences helps suppress interferences from surroundings, which present dense and redundant patterns shown in the experiments, the kernel norm of interferences is helpful. And the operator is the shrinkage-threshold function on its singular values.

### 3. EXPERIMENTS

Experiments are conducted on data from two 3D radar imaging systems for the three tasks mentioned earlier. The main parameters are listed in Tab.1, and the system’s photos are shown in Fig.3.

#### 3.1. Task 1: scattering diagnosis task

The scene shown in Fig.4(a) contains three scatters of different radar cross sections (RCS). Due to limited pages, we just present 2D projected top-view images in this one. The sparse-oriented is based on \( l_1 \) norm. We can see that compared with the MF result in Fig.4(b), ours in Fig.4(d) has less clutter and noise, and the resolution is higher. And compared with the sparse-oriented (SRO) result in Fig.4(c), ours has higher energy precision. This is due to the task demands being taken into consideration, with the \( l_q \) norm to avoid biased estimation. Qualitatively, the relative energy errors of three scatters are calculated for the last two. The results of SRO are 8.4%, 14.2%, and 37.3%, respectively. While the proposed one, they are 2.9%, 3.8%, and 1.2%, respectively. The average error is 2.6%, with an increment of 17.4%.

#### 3.2. Task 2: person screen imaging

The scene in Fig. 5(a) depicts a rifle model in free space, without the rifle being placed on a human body model to minimize the impact of thin clothing on the target’s imaging result. In Fig. 5(b), the result of MF shows a significant amount of clutter and noise. SRO’s result in Fig. 5(c) exhibits a severe loss of shape information. Our proposed method in Fig. 5(d) achieves the best visual performance, striking a balance between the previous two approaches. To evaluate the shape information loss quantitatively, we calculate the SSIM metric. The reference image is the

![Experiment systems. The left one is for task 1, and the right one is for task 2 and task 3.](image)

---

### Table 1: Experiment Parameters

| Task  | Center frequency | Bandwidth | Array size | Center range |
|-------|------------------|-----------|------------|--------------|
| Task 1 | 11GHz            | 2GHz      | 5m x 5m   | \( \approx 15 \)m |
| Task 2 | 79GHz            | 4GHz      | 0.4m x 0.4m | \( \approx 0.6 \)m |
| Task 3 | 79GHz            | 4GHz      | 0.4m x 0.4m | \( \approx 0.6 \)m |
MF image enhanced through post-processing with hand-crafted enhancements and the target-extraction algorithm. The SSIM scores are 0.51 (MF), 0.79 (SRO), and 0.95 (proposed). A higher score indicates a better retrieval of shape information. Our method achieves the highest score, aligning with the visual perception.

### 3.3. Task 3: parcel screen imaging task

The scene in Fig. 6(a) shows a rifle model in a box. The box's influence cannot be ignored. The MF result in Fig. 6(b) reveals that the rifle image is covered by the box's interference, which has a redundant and dense pattern. Mathematically, it is low-rank and characterized using the kernel norm. In Fig. 6(d), our proposed framework successfully suppresses the interference while preserving the shape to some extent. In comparison, the SRO result in Fig. 6(c) is much worse as it only retrieves a few strong scatters, leading to ineffective target detection and recognition. Our result exhibits structure loss due to interference. Quantitatively, we evaluate the suppression performance using the target-to-background ratio (TBR) metric, yielding 19.5dB (MF) and 56.2dB (proposed) with a 36.7dB suppression gain. These experiments demonstrate the effectiveness of our approach, showing its flexibility in handling various tasks and achieving impressive performance in task-oriented information retrieval.

### 4. CONCLUSION

To address the 3D radar imaging inverse problems, we introduce a novel multi-cognition task-oriented imaging framework. With the advancement of wide-bandwidth radar systems, there is an opportunity to retrieve more information. Our goal is to enable task-specific information retrieval, a topic that has been largely overlooked in this field. Unlike existing approaches that mainly differentiate between the target and background, we go a step further by analyzing and distinguishing diverse task demands. These demands are then addressed using a multi-cognition regularization model. We propose a generalized imaging method, leveraging approximation and variable-splitting techniques, to solve this model. Experimental results across various tasks demonstrate the effectiveness of our framework.

### 5. REFERENCES

[1] K. Qian, Y. Wang, Y. Shi, and X. X. Zhu, “γ-Net: Superresolving SAR Tomographic Inversion via Deep Learning,” IEEE Trans. Geosci. Remote Sensing, vol. 60, pp. 1–16, 2022, doi: 10.1109/TGRS.2022.3164193.

[2] Z. Jiao et al., “Urban 3D imaging using airborne TomoSAR: Contextual information-based approach in the statistical way,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 170, pp. 127–141, Dec. 2020, doi: 10.1016/j.isprsjprs.2020.10.013.

[3] J. Alvarez, “Near-Field 2-D-Lateral Scan System for RCS Measurement of Full-Scale Targets Located on the Ground,” IEEE Trans. Antennas Propagat., vol. 67, no. 6, pp. 4049–4058, Jun. 2019, doi: 10.1109/TAP.2019.2905727.

[4] L. Pu, X. Zhang, J. Shi, B. Tian, and S. Wei, “Sparse Bayesian 3-D Imaging for Low-RCS Objects via Dyadic Green’s Function,” Antennas Wirel. Propag. Lett., vol. 20, no. 8, pp. 1537–1541, Aug. 2021, doi: 10.1109/LAWP.2021.3090052.

[5] S. Wei et al., “3DRIED: A High-Resolution 3-D Millimeter-Wave Radar Dataset Dedicated to Imaging and Evaluation,” Remote Sensing, vol. 13, no. 17, p. 3366, Aug. 2021, doi: 10.3390/rs13173366.