Progress In Electromagnetics Research Letters, Vol. 55, 15–22, 2015

Time Domain Sparse Representation for Multi-Aspect SAR Data of Targets

Jinrong Zhong*, Gongjian Wen, Conghui Ma, and Baiyuan Ding

Abstract—Sparse representation is the fundamental technology of compressive sensing, sparse three-dimensional (3-D) imaging, and dictionary-based parameter estimation. Typical sparse representation models of radar signal work in the frequency domain, which may encounter high dimension and large data amount of dictionary. This paper presents a time-domain (TD) representation model for multi-aspect SAR data. We generate the multi-aspect two-dimensional (2-D) TD responses of the 3-D scattering center model. Then we cut off the low-energy area of the 2-D TD response and use cutoff responses to construct the dictionary of sparse representation. Such a TD dictionary is a sparse matrix. Moreover, we build and solve the sparse representation model based on the TD dictionary. Compared with the frequency-domain (FD) sparse representation model, the data size of our TD dictionary is remarkably lower, and the solving of TD sparse representation problem is in higher efficiency. We utilize the TD sparse representation to reconstruct 3-D images from multi-aspect SAR data. Experimental results demonstrate the effectiveness and efficiency of the TD sparse representation model.

1. INTRODUCTION

Sparse representation of radar signals [1] is the fundamental for many other techniques, such as compressed sensing [2], sparse 3-D imaging [3], data compression [4], and dictionary-based parameter estimation [5]. Multi-aspect SAR data are a typical collecting geometry of radar [6]. Sparse representation of multi-aspect SAR data is of significance.

The existing sparse representation models [1–5] for radar signal are established in the frequency domain. In other words, these sparse representation methods built the dictionary and solved the model in the frequency domain. Based on frequency-domain (FD) sparse representation model, many achievements [1–5, 7, 8] have been made in the files of 3-D imaging and feature extraction. However, the larger data amount of the dictionary, accompanied with a huge computation, is always a significant challenge.

For the sparse representation problem of multi-aspect SAR data, this paper proposes a time-domain (TD) model, which constructs the dictionary in the time domain and solves the sparse representation problem in time domain too. We use 2-D time responses of the 3-D scattering center model at every aspect to construct the dictionary. Since the energy of 2-D TD responses concentrates on several small regions, we cut off the low-energy areas to make the TD dictionary become a sparse matrix. Moreover, the signal to noise rate (SNR) of the valid regions increases in the time domain. So the time domain sparse representation reduces the data amount of dictionaries substantially and improves the searching efficiency. Apply the proposed TD sparse representation to the multi-aspect SAR data of 3-D imaging. The experimental results demonstrate that TD sparse representation is effective and exhibit its efficiency.

Received 24 April 2015, Accepted 14 July 2015, Scheduled 14 August 2015
* Corresponding author: Jinrong Zhong (hytler0110@163.com).
The authors are with the ATR Laboratory, National University of Defense Technology, Changsha 410073, China.

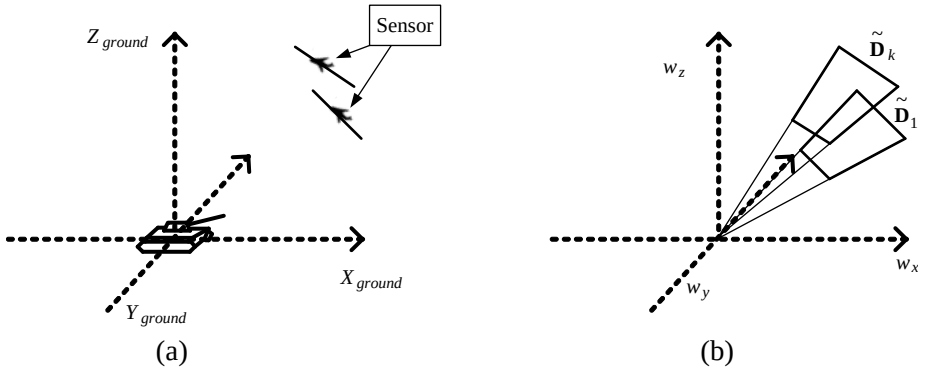

Figure 1. Multiple-aspect SAR collecting geometry. (a) Collecting geometry. (b) Fractions in wavenumber coordinate system.

2. MULTI-ASPECT COLLATED SAR DATA

The multi-aspect collected geometry is depicted in Figure 1. At the observed angle (ϕ, θ), the radar received signal is $r(t;\theta,\phi) = [\iiint_\Lambda g(x,y,z;\theta,\phi)dxdydz] * s(t)$, and $s(t)$ is the sent signal, $*$ the convolution. Using the signal model presented by Austin et al. in [3], interpret multi-aspect SAR data in the 3-D Fourier transformation $G(w_x,w_y,w_z)$ of the reflectivity function $g(x,y,z;\phi,\theta)$. As can be seen below, our method does nothing matter with the polarization, and we take single polarization in consideration only. The signal model in Fourier transformation is given by

$$G(w_x,w_y,w_z) = \sum_{n=1}^{N} g(x,y,z;\phi,\theta)\cdot e^{-j(x_n w_x + y_n w_y + z_n w_z)} \tag{1}$$

Suppose that the center frequency of k-th aspect is $f_{c,k}$, the bandwidth B_k, and the aperture $\bar{\phi}_k$, $k = 1,\ldots,K$. By the projection-slice theorem, each aperture is a fraction in $G(w_x,w_y,w_z)$. Denote the k-th fraction's sample point set as $M_k = \{(f_j,\phi_i)\}_{j=1,i=1}^{Nf_k,Mf_k}$. The coordinates in wavenumber domain are $w_x^{j,i,k} = -4\pi f_j \cos\theta_{i,k}\cos\phi_{i,k}/c$, $w_y^{j,i,k} = -4\pi f_j \cos\theta_{i,k}\sin\phi_{i,k}/c$, $w_z^{j,i,k} = -4\pi f_j \sin\theta_{i,k}/c$. Reshape them into a vector $\mathbf{H}_M = \cup_{k=1}^{K}\mathbf{H}_{M,k} = \{\hbar_m\}_{m=1}^{M}$, $M = \sum_{k=1}^{K} M_k$, and the m-th coordinate is $\hbar_m = (w_{x,m}, w_{y,m}, w_{z,m})$.

Denote the multi-aspect SAR dataset as $\tilde{D} = \{\tilde{\mathbf{D}}_k\}_{k=1}^{K}$, and $\tilde{\mathbf{D}}_k = [G(w_x^{j,i,k}, w_y^{j,i,k}, w_z^{j,i,k})]_{i=1,j=1}^{Mf_k,Nf_k}$ is the k aspect fraction. Transform $\tilde{\mathbf{D}}_k$ into a 2-D SAR complex image and denote it as $\mathbf{D}_k \in C^{Mz_k\times Nz_k}$, which can be viewed as 2-D time domain data. Mz_k and Nz_k are the row and column numbers.

3. FREQUENCY DOMAIN SPARSE REPRESENTATION

Sample 3-D imaging space $\mathbf{Q}$ into a series of 3-D locations as candidate locations,

$$\mathbf{Q}_M = \{\lambda_1,\ldots,\lambda_N\} = \{(x_n,y_n,z_n)\}_{n=1}^{N} \tag{2}$$

Usually $\mathbf{Q}_M$ is chosen on a uniform rectilinear grid. The grid point number is N. Take the isotropic point scattering center (IPS) model to construct the dictionary and sparse representative the target. The dictionary is built in the frequency domain, using the FD responses of IPS

$$\tilde{\mathbf{Y}} = F(\mathbf{H}_M,\mathbf{Q}_M) = \left[e^{-j(x_n w_{x,m} + y_n w_{y,m} + z_n w_{z,m})}\right]_{m,n} \tag{3}$$

where m indexes the M measured frequency (or wavenumber) domain down rows, and n indexes the location across columns. Based on $\tilde{\mathbf{Y}}$, the vectored measured data can be approximated as

$$\tilde{\mathbf{d}} = \tilde{\mathbf{Y}}\cdot\mathbf{s} + \tilde{\mathbf{v}} \tag{4}$$

where $\tilde{\mathbf{d}} \in \mathcal{C}^{M\times 1}$ is vectored from the multi-aspect SAR data, and $\tilde{\mathbf{v}}$ is the noise. $\mathbf{s}$ is a sparse N-dimensional vector and also the scattering amplitude vector. $\hat{\mathbf{s}}$ is the solution of the sparse optimization problem

$$\hat{\mathbf{s}} = \arg\min \|\mathbf{s}\|_0 \ \text{ s.t. } \left\|\tilde{\mathbf{Y}}\mathbf{s} - \mathbf{d}\right\|_2 < \varepsilon \tag{5}$$

4. TIME DOMAIN SPARSE REPRESENTATION

The 3-D scattering center models have a natural characteristic: the amplitudes of its 2-D response at every aspect are well distributed in the frequency domain while the amplitudes are uneven in the time domain. We develop a TD sparse representation model, exploiting such a property.

As presented above, for a location $\lambda_n = [x_n, y_n, z_n]^T$ in $\mathbf{Q}_M$, the 2-D FD response at the k aspect is

$$\tilde{E}(\mathbf{H}_{M,k}|\lambda_n) = \left[\exp\left[-j\left(x_n w_{x,k}^{i,j} + y_n w_{y,k}^{i,j} + z_n w_{z,k}^{i,j}\right)\right]\right]_{i=1,j=1}^{Mf_k, Nf_k} \tag{6}$$

$\tilde{\mathbf{E}}_{k,n} = \tilde{E}(\mathbf{H}_{M,k}|\lambda_n)$ is a $Mf_k \times Nf_k$ matrix. Interpolate and transform $\tilde{\mathbf{E}}_n$ to the time domain

$$\mathbf{E}_{k,n} = F_2^{-1} \cdot B \cdot \tilde{\mathbf{E}}_{k,n} \tag{7}$$

Here, $F_2^{-1}(\cdot)$ represents the 2-D inverse fast Fourier transform (IFFT), and B represents interpolate the 2-D sector data into rectangular data. Figure 2 shows an example of the 2-D frequency/time domain responses at the k aspect. It can be found that the energy of TD response, $\mathbf{E}_{k,n} = [e_{i,j,k}]_{i=1,j=1}^{Mz_k \times Nz_k}$, concentrates in a small region. The other big region is lower than -50 dB of the peak-power. This characteristic does not depend on the measurement data, but a natural characteristic scattering center model. We cut off the small amplitude regions

$$e_{i,j,k} = \begin{cases} 0 & |e_{i,j,k}| < \xi_n \\ e_{i,j,k} & \text{else} \end{cases} \tag{8}$$

Here, ξ_n is an adaptive threshold. Its value is determined by equation $\xi_n = \max(|\mathbf{E}_{k,n}|) \cdot 10^{\beta_d/20}$. $\max(|\mathbf{E}_{k,n}|)$ is the maximum amplitude of $\mathbf{E}_{k,n}$. $\beta_d (\beta_d < 0)$ is an experiential threshold for the whole dictionary. Write such processing in the form of an operator, $\mathbf{E}_{k,n}^* = \Gamma^{\beta_d}(\mathbf{E}_{k,n})$. We stack the cutoff 2-D response $\mathbf{E}_{k,n}^*$ into a vector denote as $\boldsymbol{\psi}_{k,n} = \text{vec}(\mathbf{E}_{k,n}^*) = [\varphi_{1,n}, \ldots, \varphi_{M_k,n}]^T$. It is the component of λ_n at the k-th aspect. We express the process of getting $\boldsymbol{\psi}_{k,n}$ in the form of operators, $\boldsymbol{\psi}_{k,n} = \text{vec} \cdot \Gamma^{\beta_d} \cdot F_2^{-1} \cdot B \cdot \tilde{\mathbf{E}}_{k,n}$. The k part of the TD dictionary, $\mathbf{Y}_k$, consists of all the $\boldsymbol{\psi}_{k,n}$ at k aspects for all λ_n, $n = 1, \ldots, N$. After obtaining all the K parts, we construct the dictionary of $\mathbf{Q}_M$ for the multi-aspect SAR data as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_K \end{bmatrix} \tag{9}$$

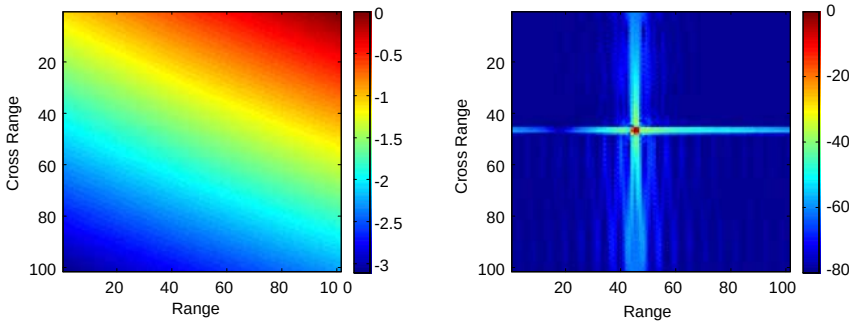

Figure 2. 2-D frequency/time domain responses of 3-D scattering center model at someone aspect.

When $\beta_d > -\infty$, $\mathbf{Y}$ is a sparse 2-D matrix. Suppose that Q is the number of nonzero elements and that $\eta = Q/(M \cdot N)$ is the proportion of nonzero elements and indicates the sparse degree of the dictionary matrix. The larger β_d is, the smaller η is, and the sparser is $\mathbf{Y}$, which is beneficial to reducing the data size of dictionaries. In addition, $\mathbf{Y}$ being a sparse matrix, is beneficial to reducing the burden of computation.

Interpolate and transform $\tilde{\mathbf{D}}_k$ to the time domain, and then stack it into a M_k-dimensional vector $\mathbf{d}_k = [d, \ldots, d_m, \ldots, d_{M_k}]^T = \text{vec} \cdot F_2^{-1} \cdot B \cdot \tilde{\mathbf{D}}_k$. So, the measured data can be approximated as

$$\mathbf{d} = \mathbf{Y}\mathbf{s} + \mathbf{v}, \quad \text{here}, \quad \mathbf{d} = \begin{bmatrix} \mathbf{d}_1 \\ \vdots \\ \mathbf{d}_{N_{ap}} \end{bmatrix} \tag{10}$$

where $\mathbf{v}$ is TD noise vector, $\mathbf{s}$ a sparse vector, and $\mathbf{d}$ the vectored multi-aspect SAR TD data. According to the time-frequency duality principle, as $\beta_d = -\infty$ (in other words, the time-domain dictionary is not cut off), the FD model (4) and TD model (10) are equivalent. Also, their solutions should be the same. As $\beta_d \geq -\infty$, due to truncation processing, which make the dictionary loss energy and rice the atomic correlation, the solutions of the model (4) and model (10) may be different. But it cannot assert that the convergence and accuracy of model (10) is inferior to model (4). Because the energy of target responses congregates in TD, the SNR rises in the active area. And the truncation not only cut off signals in the dictionary, but also cut off the noise in the measurement data covertly. These factors are conducive to the solution. Solve the sparse optimization problem to obtain $\hat{\mathbf{s}}^*$,

$$\hat{\mathbf{s}}^* = \arg\min \|\mathbf{s}\|_0 \ \text{ s.t. } \ \|\mathbf{Y}\mathbf{s} - \mathbf{d}\|_2^2 < \varepsilon \tag{11}$$

Since l_0-regulated optimization is hard to solve, we take the l_1-regulated technique [9, 11] to deal with problem (11)

$$\hat{\mathbf{s}}^* = \arg\min \left\{ \|\mathbf{Y}\mathbf{s} - \mathbf{d}\|_2^2 + \lambda \|\mathbf{s}\|_1 \right\} \tag{12}$$

5. EXPERIMENTAL RESULTS

We use the TD sparse model to representative multi-aspect SAR data simulated by GTD [10] model, then compare the performances of TD model with the FD model. The performances include validity and efficiency. In order to demonstrate the validity of the TD sparse representation, we compare the sparse solutions, scattering center locations and the reconstructed 2-D SAR images with the FD method. In order to demonstrate the efficiency, we compare their dictionaries' data size and time consumption of solving.

The simulated complex target consists of 11 scattering centers, which are located in range $\mathbf{Q} = \{x, y, z | -2 < x < 2, -2 < y < 2, 0 < z < 1\}$. The locations of them are listed in Table 1. Use GTD-based scattering center model [10] to produce multiple-aspect SAR data

$$G(w_x, w_y, w_z) = \sum_{i=1}^{P} A_i \cdot (jf/f_c)^{\alpha_i} \cdot e^{-j(x_i w_x + y_i w_y + z_i w_z)} + v$$

Table 1. The locations of target's 11 scattering centers.

Index	$[x, y, z]$	Index	$[x, y, z]$
1	[−1.727, 0.890, 0.771]	**7**	[1.598, −0.499, 0.512]
2	[0.975, 0.750, 0.635]	**8**	[−1.727, 0.250, 0.771]
3	[0.975, 0.499, 0.635]	**9**	[1.598, 0.8895, 0.512]
4	[0.975, −0.499, 0.635]	**10**	[−1.727, −0.250, 0.771]
5	[−1.727, −0.890, 0.771]	**11**	[0.499, 0.044, 0.685]
6	[1.598, 0.499, 0.512]	**12**	

A_i is randomly generated, and frequency-dependent factor α_i is set to one of $[-1, -1/2, 0, 1/2, 1]$ randomly. Produce SAR data at five aspects (θ_k, ϕ_k). The azimuths are $\phi_k = 5°$, $k = 1, \ldots, 5$, and the elevations are $\theta_1 = 22.5°$, $\theta_2 = 23.5°$, $\theta_3 = 24.5°$, $\theta_4 = 25.5°$, $\theta_5 = 27.5°$. The sample intervals are $\Delta f = 30\,\text{MHz}$ and $\Delta\phi = 0.1°$. Add zero-mean white Gaussian noise $v \sim N(0, \sigma_v^2)$ to make the FD $SNR_{Fd} = -10\,\text{dB}$, TD $SNR_{img} = 20\,\text{dB}$. The five SAR images with noise are shown in Figure 3, and the dynamic range for display are $[-40,\ 0]\,\text{dB}$.

Take intervals $\Delta x = 0.1\,\text{m}$, $\Delta y = 0.1\,\text{m}$, $\Delta z = 0.1\,\text{m}$ to grid the 3-D space $\mathbf{Q}$ into a set of 3-D locations, $N = 18491$ locations in total. So, the sparse dictionaries' demotion is $M \times N = 4420 \times 18491$. Construct the TD dictionary with a threshold $\beta_d = -50\,\text{dB}$. The proportion of nonzero elements is $\eta = 7.1\%$. It is lower than a FD dictionary in an order of magnitude. Solve the sparse representation problem $\tilde{\mathbf{d}} = \tilde{\mathbf{Y}} \cdot \mathbf{s} + \tilde{\mathbf{v}}$ and $\mathbf{d} = \mathbf{Y} \cdot \mathbf{s} + \mathbf{v}$. The sparse vectors $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}^*$ are shown in Figure 4. Their corresponding locations are shown in Figure 5. As an example, we use $\hat{\mathbf{s}}^*$ to reconstruct the SAR image at $(\theta_1 = 25.5°, \phi_k = 5°)$ by $\mathbf{Y} \cdot \hat{\mathbf{s}}^*$ directly, then use $\hat{\mathbf{s}}$ to rebuild the FD SAR data and turn it into image-domain. The reconstructed images are shown in Figure 6.

From Figure 4, it can be seen that $\hat{\mathbf{s}}^*$ is convergent to the real sparse vector. There is little difference between sparse vectors $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}^*$. Figure 5 confirms such a judgment. Since the scattering centers are not

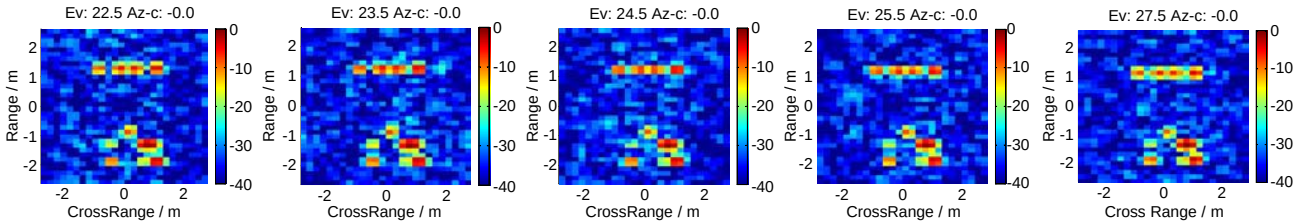

Figure 3. The simulated five-aspect SAR images.

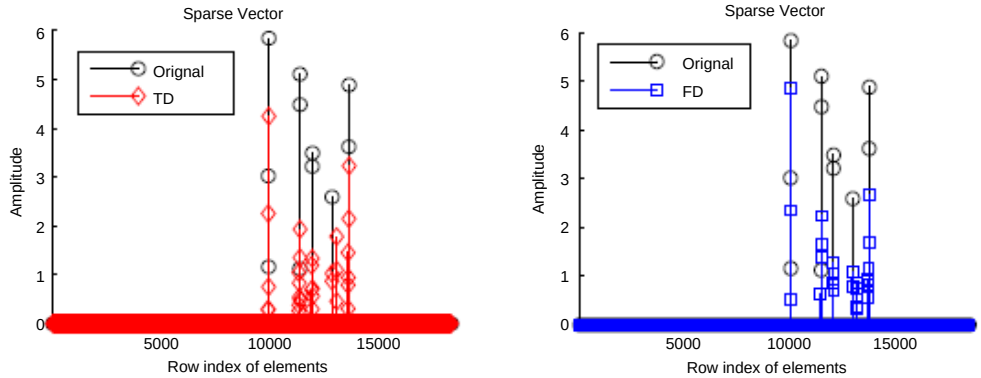

Figure 4. Sparse vectors $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}^*$.

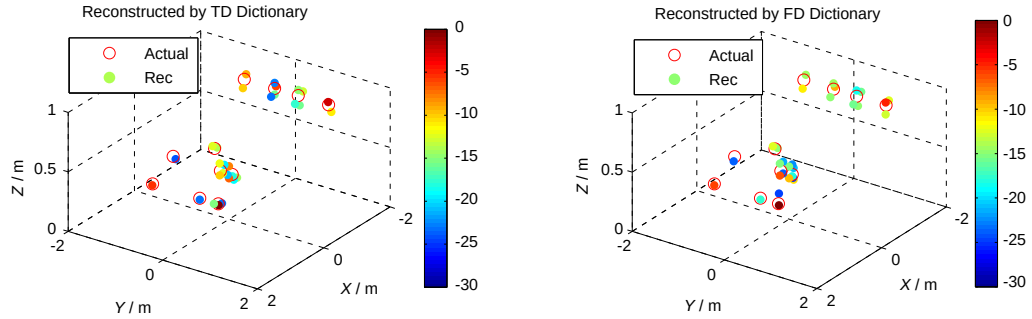

Figure 5. The corresponding 3-D locations of the nonzero elements.

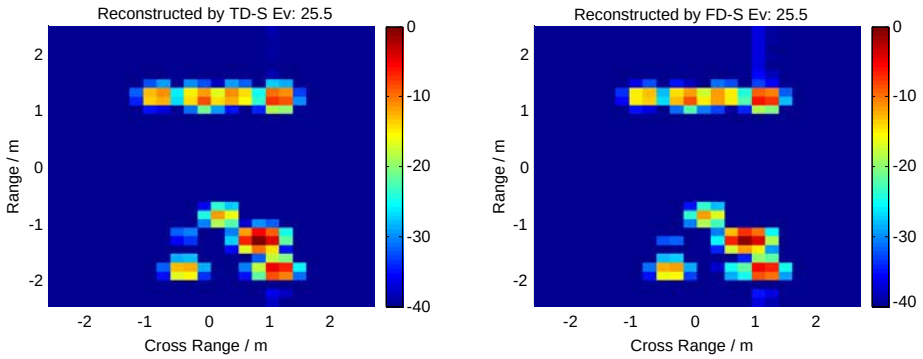

Figure 6. The reconstructed SAR images.

right on the grid, the nonzero elements are located around the real scattering centers. The color is the value of the nonzero elements, also is the amplitude locations. They are can be viewed as 3-D images of the simulated target. We can see that the two almost identical images are reasonable. Both the TD and FD dictionary sparse representations can availably represent the target. In Figure 6, the reconstructed SAR images are with less noise. The first one is reconstructed by TD sparse representation, and it is 0.958 similar to the original noise-free SAR complex image. The second one is reconstructed by FD sparse representation, and its similarity is 0.958. The formula to calculate the similarity is

$$\mu = (\hat{\mathbf{e}} - \mathbf{e})^H \cdot (\hat{\mathbf{e}} - \mathbf{e}) \,/\, \left(\mathbf{e}^H \cdot \mathbf{e}\right)$$

From the results presented above, it can be said that the TD sparse representation is efficacious.

Below, we discuss the efficiency. First, as mentioned above, compared with the FD dictionary, the data size of our TD dictionary reduces 92.9%. Second, we examine the computation complexity of solving our sparse representation problem, in comparison with the FD one. In this experiment, we use the method proposed by Kim et al. for l_1-regularized least squares in [11] to solve the TD and FD sparse representation problems. The corresponding $l_1\_l_s$ Matlab solver is presented in [12]. The TD sparse representation costs 62.3 seconds in total to obtain the solution. In detail, it needs 46 iterations, and each iteration has to take 11.7 preconditioned conjugate gradient (PCG) iterations in average. The FD sparse representation requires 72 iterations, and each iteration has to take 34.5 PCG iterations in average. It costs 843.4 seconds in total. Even if we do not cut off the TD dictionary, set $\beta_d = -\infty$, the two models are perfectly equivalent. The TD sparse representation costs 376.3 seconds, which is less than the FD model. The large number of elements in the TD dictionary contribute to the increase in efficiency. One reason for such a contribution is that the multiplication of $\mathbf{Ys}$ requires fewer multiplies. Besides, we find that there is another reason, i.e., TD sparse representation has advantages in avoiding noise jamming. As the low-energy area of the responses in the TD dictionary are cut off, it is equivalent to removing most noise of the measured data in the solving process. The fundamental mechanism can be abstractly and briefly explained as following. For the sparse representation, we have to search the solution that makes $\Delta = ||\mathbf{d} - \mathbf{Ys}||_2^2 + \lambda||\mathbf{s}||_1$ reach its minimization. When the locations of the nonzero elements in $\mathbf{s}$ are fixed, the zero-area of $\mathbf{Ys}$, denoted as $[\mathbf{Ys}]^{\mathbf{0}}$, is the noise area determined by the optimization algorithm automatically. $||[\mathbf{d}]^0 - [\mathbf{Ys}]^{\mathbf{0}}||_2^2 = c$ is a constant, which is helpful to smoothing the cost-function, excluding the impact of noise on the optimization iterating. As the nonzero elements in $\mathbf{s}$ change, it is equivalent to adaptively redesignating the target region and noise area. After each iteration, the iterated $\mathbf{s}$ approaches the truth solution. The separation of target regions and noise area becomes more and more reasonable. Compared with denoising by directly truncating noise area in the image domain, this embedding method is advantageous. The noise is removed more accurately. It may accelerate the convergence rate of the optimization and can be demonstrated by the iterative situation in simulation results. We have noticed that the time consumption of solving the TD sparse representation problem is much less than the FD one. We use apagoge to prove that the cost function becomes smooth. We make a hypothesis that the only reason of the increase in efficiency is that less multiplies are needed. If this hypothesis is true, the time consumption of every iteration will reduce, but the iteration number will not. However, in our experiment, not only the time consumption reduces,

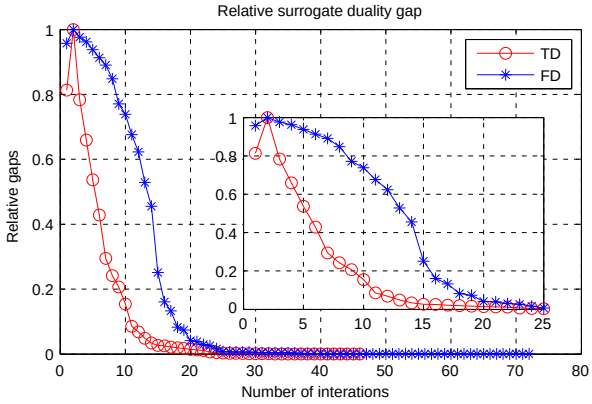

Figure 7. The relative surrogate duality gaps of TD method and FD method.

but also the iteration number decreases. We plot the relative surrogate duality gaps of TD method and FD method again in Figure 7 with an enlarged subplot. It suggest that the cost function is smoothed. Furthermore, we try to explain the mechanism in detail by theoretical analysis. Suppose that $\tilde{\mathbf{Y}}$ is the FD dictionary. As $\mathbf{d} = \mathbf{e} + \mathbf{v}$, the searching for the minimum will be interfered noise $\mathbf{v}$.

$$\Delta(\mathbf{s}) = \left\|\mathbf{d} - \tilde{\mathbf{Y}}\mathbf{s}\right\|_2^2 + \lambda||\mathbf{s}||_1$$

We index the iteration times with it and suppose that the desired search direction at current point $\mathbf{s}_{it}$ is $\vec{\nabla}_{it}^0$

$$\Delta^0\left(\mathbf{s}_{it} + \vec{\nabla}_{it}^0\right) = \left\|\mathbf{e} - \tilde{\mathbf{Y}}\left(\mathbf{s}_{it} + \vec{\nabla}_{it}^0\right)\right\|_2^2 + \lambda\left\|\mathbf{s}_{it} + \vec{\nabla}_{it}^0\right\|_1$$

Due to the noise, there may be another disturbance direction $\vec{\nabla}_{it}^*$ that makes $\Delta(\mathbf{s}_{it} + \vec{\nabla}_{it}^*) \leq \Delta(\mathbf{s}_{it} + \vec{\nabla}_{it}^0)$. This is probabilistic. The larger is the noise, the greater is the probability. In our TD model, we transform $\tilde{\mathbf{Y}}$ into TD first. $\mathbf{Y}^*$ is the un-cutoff time-domain dictionary and $\Delta(\mathbf{s}_{it}) = ||\mathbf{d} - \mathbf{Y}^*\mathbf{s}_{it}||_2^2 + \lambda||\mathbf{s}_{it}||_1$. Suppose that $[\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L$ is the low energy area of $\mathbf{Y}^*\mathbf{s}_{it}$, in which the elements will be set to zero. $[\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma$ is the high energy area, in which the elements maintain the original values. $[\cdot]_{it}^L \cap [\cdot]_{it}^\Gamma = \emptyset$. We rewrite $\Delta_1(\mathbf{s}_{it}) = ||\mathbf{d} - \mathbf{Y}^*\mathbf{s}_{it}||_2^2$ as,

$$\begin{aligned}\Delta_1 = \|\mathbf{d} - \mathbf{Y}^*\mathbf{s}_{it}\|_2^2 &= \left\|[\mathbf{d}]_{it}^L + [\mathbf{d}_{it}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma\right\|_2 \\ &= \left\|[\mathbf{d}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L\right\|_2^2 + \left\|[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma\right\|_2^2 + 2\left([\mathbf{d}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L\right)^H\left([\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma\right)\end{aligned}$$

Here, $[\cdot]_{it}^L$ and $[\cdot]_{it}^\Gamma$ are determinate according to $\mathbf{Y}^*\mathbf{s}_{it}$ adaptively. $[\mathbf{d}]_{it}^L$ represents that the elements in $\mathbf{d}$ corresponding to the low-energy area $[\cdot]_{it}^L$ remain, and the rest are set to zero. $[\mathbf{d}]_{it}^\Gamma$ represents that the elements in $\mathbf{d}$ corresponding to the area $[\cdot]_{it}^\Gamma$. Since $[\cdot]_{it}^L \cap [\cdot]_{it}^\Gamma = \emptyset$, $([\mathbf{d}]_{it}^{\mathbf{0}} - [\mathbf{Y}\mathbf{s}]_{it}^{\mathbf{0}})^H([\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}\mathbf{s}]_{it}^\Gamma) = 0$, so $\Delta_1(\mathbf{s}_{it}) = ||[\mathbf{d}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L||_2^2 + ||[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma||_2^2$. Then,

$$\Delta(\mathbf{s}_{it}) = \left\|[\mathbf{d}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L\right\|_2^2 + \left\|[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma\right\|_2^2 + \lambda\left\|\mathbf{s}_{it}\right\|_1$$

The SNR of $[\mathbf{d}]_{it}^\Gamma$ is higher than the SNR of $[\mathbf{d}]_{it}^L$, and the convergence of $||[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma||_2^2$ should be better than $||[\mathbf{d}]_{it}^L - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^L||_2^2$. In our TD method, we set $[\mathbf{Y}\mathbf{s}]_{it}^L = \mathbf{0}$. $[\cdot]_{it}^L$ can also be expressed as $[\cdot]_{it}^{\mathbf{0}}$.

$$\Delta(\mathbf{s}_{it}) = \left\|[\mathbf{d}]_{it}^{\mathbf{0}}\right\|_2^2 + \left\|[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma\right\|_2^2 + \lambda\left\|\mathbf{s}_{it}\right\|_1$$

If $[\cdot]_{it+1}^{\mathbf{0}}$ is not changed, $||[\mathbf{d}]_{it}^{\mathbf{0}}||_2^2 = ||[\mathbf{d}]_{it+1}^{\mathbf{0}}||_2^2$ is a constant, which will not affect the searching of descent direction. If $[\cdot]_{it+1}^{\mathbf{0}}$ is changed, for a right descent direction $\vec{\nabla}_{it}$, $||[\mathbf{d}]_{it}^{\mathbf{0}}||_2^2 \leq [\mathbf{d}]_{it+1}^{\mathbf{0}}||_2^2$ must be true. The convergence of $||[\mathbf{d}]_{it}^{\mathbf{0}}||_2^2$ keeps synchronous with $||[\mathbf{d}]_{it}^\Gamma - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^\Gamma||_2^2$. When $\mathbf{s}_{it}$ is far away from

the real solution $\mathbf{s}$, it is important and benefitial to the convergence of $\Delta(\mathbf{s})$. Of course, when $\mathbf{s}_{it}$ has been very close to the real solution $\mathbf{s}$, the convergence speed of $|||[\mathbf{d}]_{it}^{\Gamma} - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^{\Gamma}||_2^2$ will be slower than the original $||\mathbf{d} - \mathbf{Y}^*\mathbf{s}_{it}||_2^2$. Because at such case, $|||[\mathbf{d}]_{it}^{\mathbf{0}} - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^{\mathbf{0}}||_2^2$ will also converge to a minimum $\mathbf{s}$. In other words, the descent directions of $|||[\mathbf{d}]_{it}^{\mathbf{0}} - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^{\mathbf{0}}||_2^2$ and $|||[\mathbf{d}]_{it}^{\Gamma} - [\mathbf{Y}^*\mathbf{s}_{it}]_{it}^{\Gamma}||_2^2$ may be the same with great probability, and $|||[\mathbf{d}]_{it+1}^{\mathbf{0}} - [\mathbf{Y}^*\mathbf{s}_{it+1}]_{it+1}^{\mathbf{0}}||_2^2 \leq |[\mathbf{d}]_{it}^{\mathbf{0}}|_2^2$. So, it can be forecasted that the convergence speed of our TD model is slower than the FD one when $\mathbf{s}_{it}$ is close to $\mathbf{s}$ since some signal is lost as we set $[\mathbf{Y}^*\mathbf{s}_{it}]_{it}^{\mathbf{0}} = \mathbf{0}$. However, at the beginning of the iteration, $\mathbf{s}_{it}$ is far from $\mathbf{s}$, and the convergence speed of our TD model is faster. The descent directions keep the same as the noiseless case $||\mathbf{e} - \tilde{\mathbf{Y}}\mathbf{s}_{it}||_2^2 + \lambda||\mathbf{s}_{it}||_1$ with great probability when the effect of most of the noise is dispelled. As a comprehensive effect, the convergence speed and time consumption of our TD model are superior to the FD one, which has been exhibited in our experiment. When $\beta_d = -\infty$, there is no truncation, but a large number of elements in the TD dictionary approach zero. Therefore, the TD sparse representation also has such a character.

6. CONCLUSIONS

This paper proposes a concept of TD sparse representation for multi-aspect SAR data and presents the method of constructing TD dictionary and solving TD sparse representation problem. Compared with the FD model, the data size of the TD dictionary consisting of truncated 2-D TD responses is far smaller. And it improves the convergence speed of the optimization algorithm. Experimental results show that the TD sparse representation is available, and its effectiveness is considerably advanced.

REFERENCES

1. Samadi, S., M. Çetin, and M. A. Masnadi-Shirazi, "Sparse representation-based synthetic aperture radar imaging," *IET Radar, Sonar, Nav.*, Vol. 5, No. 2, 182–193, 2011.
2. Herman, M. A. and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, Vol. 57, No. 6, 2275–2284, 2009.
3. Austin, C. D., E. Ertin, and R. L. Moses, "Sparse signal methods for 3D radar imaging," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 5, No. 3, 408–423, 2011.
4. Bhattacharya, S., T. R. Blumensath, B. Mulgrew, and M. Davies, "Synthetic aperture radar raw data encoding using compressed sensing," *Proc. IEEE 2008 Int. Conf. Radar*, 1–5, 2008.
5. Austin, C. D., J. N. Ash, and R. L. Moses, "Dynamic dictionary algorithms for model order and parameter estimation," *IEEE Trans. Signal Process.*, Vol. 61, No. 20, 5117–5130, 2013.
6. Rigling, B. D. and R. L. Moses, "Three-dimensional surface reconstruction from multi-static SAR images," *IEEE Trans. Image Process.*, Vol. 14, No. 8, 1159–1171, 2003.
7. Corporation, H. P., "Sparse representation denoising for radar high-resolution range profiling," *Int. J. Antennas Propagat.*, Vol. 13, No. 1, 479–482, 2014.
8. Zhu, S., A. Mohammad-Djafari, H. Wang, B. Deng, L. Xiang, and J. Mao, "Parameter estimation for SAR micromotion target based on sparse signal representation," *EURASIP J. Advances Signal Process.*, Vol. 13, No. 2, 26–26, 2012.
9. Zhu, X. X. and R. Bamler, "Tomographic SAR inversion by L_1-norm regularization: The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, Vol. 48, No. 10, 3839–3846, 2010.
10. Potter, L. C., D. M. Chiang, R. Carriere, and J. Gerry, "A GTD-based parametric model for radar scattering," *IEEE Trans. Antennas Propagat.*, Vol. 43, No. 11, 1058–1067, 1995.
11. Kim, S. J., K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "A method for large-scale l_1-regularized least squares," *IEEE Journal of Selected Topics in Signal Processing*, Vol. 1, No. 4, 606–617, 2007.
12. http://web.stanford.edu/∼boyd/l1_ls/.