Three-Dimensional SAR Focusing via Compressive Sensing: The Case Study of Angel Stadium
Sun Xilong, Yu Anxi, Dong Zhen, and Liang Diannong

Abstract—Recently, a synthetic aperture radar (SAR) tomographic focusing method based on compressive sensing was proposed. This focusing method can reduce the required number of measurements and achieve satisfying elevation resolving ability. First, we briefly review this novel focusing method and prove the applicability of compressed sensing (CS) for SAR tomography theoretically using the latest improvement of CS. Then, we apply this focusing method to the 3-D reconstruction of Angel Stadium with Envisat-ASAR data. Both the theoretical analysis and satisfying results of real data processing confirmed the applicability of this SAR tomographic focusing method.

Index Terms—Compressive sensing (CS), Envisat-ASAR, synthetic aperture radar (SAR), tomography.

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) tomography forms a synthetic aperture along the normal-slant-range (nsr) direction to achieve resolving ability of multiple scatterers within one range-azimuth resolution cell [1], [2]. SAR tomography enables SAR to have the ability of 3-D imaging, so it has attracted an explosion of interest in the last few years. At present, the synthetic aperture along nsr direction of SAR tomography is usually achieved by repeat-pass missions of single-antenna SAR systems. In this data acquisition mode, particularly for spaceborne SAR, the passes are unevenly spaced, and available passes are much lower than that needed for a large synthetic aperture. Hence, the urgent problem of SAR tomography is how to achieve satisfying elevation resolution with a small number of nonuniform passes.

Recently, a SAR tomographic focusing method based on compressive sensing, in which the inverse problem of SAR tomography is solved by $\ell_1$-norm regularization, was proposed in [3]–[5]. This SAR tomographic focusing method can reduce the required number of measurements and achieve satisfying elevation resolving ability with a given aperture in nsr direction. In [3] and [4], the required number of measurements was discussed, and experimental results of ERS1/2 data were presented. In [5], Super-resolution properties and point localization accuracies of this focusing method were demonstrated using simulations and real TerraSAR-X data.

In SAR tomography, the measurements are random samples of continuous frequency spectrum of the unknown signal, instead of random samples of discrete Fourier transform (DFT). As a consequence, the sensing matrix is not obtained by randomly sampling rows of an orthonormal matrix. The theory of CS in [6]–[11] cannot provide a framework for this measurement strategy perfectly. Recently, Candes introduced a simple and very general theory of compressive sensing in [12] that can provide a framework for the measurement strategy in SAR tomography. In this letter, we will prove the applicability of compressed sensing (CS) for SAR tomography theoretically using the theory in [12]. Then, we will apply the CS-based focusing method to real data to validate this method. The Envisat-ASAR data over Angel Stadium between October 29th 2003 and November 7th 2007 are used.

II. FORMULATION OF SAR TOMOGRAPHY

To achieve resolving ability of multiple scatterers within one azimuth-range resolution cell, a synthetic aperture along the nsr direction is needed to form by repeat-pass missions of single-antenna SAR system. Fig. 1 gives the multipass acquisition geometry, for the sake of simplicity, all the acquisitions are aligned along normal-slant-range direction.

Fig. 1. Multipass acquisition geometry of SAR tomography. For the sake of simplicity, all the acquisitions are aligned along normal-slant-range direction.
First, classical 2-D SAR focusing is applied to the echoes, \( M \) SAR images are acquired consequently. Then, one of the images is selected as the reference image, and each of the others is registered with reference to it. After the process of deramping, the complex values of a fixed pixel in \( m \)th image can be written as the following form [1]:

\[
g_m = \int_{-s_{\text{max}}}^{s_{\text{max}}} \gamma(s) \exp(j2\pi \xi_ms)ds, \quad m = 1, 2, \ldots, M
\]

\[
\xi_m = 2b_{\perp m}/\lambda r
\]

where \( s \) is the nsr elevation, \( 2s_{\text{max}} \) is the extent of scene along nsr direction within one azimuth-range resolution cell, \( \gamma(s) \) is the function that models the reflectivity profile along the nsr direction. \( \lambda \) is the wave length, \( r \) is the slant range from the sensor position of reference image to scene center, \( b_{\perp m} \) is the orthogonal baseline. Equation (1) shows that the complex values of a fixed pixel in acquisitions, received at different spatial positions, are samples of continuous frequency spectrum of \( \gamma(s) \) at the frequencies described by (2).

The bandwidth of acquisitions is

\[
W_{\xi} = 2B_{\perp}/\lambda r
\]

where \( B_{\perp} = b_{\perp M} - b_{\perp 1} \) is the overall orthogonal baseline length. The Rayleigh resolution of nsr elevation corresponding to the bandwidth is

\[
\rho_s = \lambda r/2B_{\perp}.
\]

The critical extent of scene, which does not induce ambiguities in the reconstructed profile, can be given by the following equation approximately:

\[
2s_{\text{max}} = \lambda r/2\Delta b_{\perp}
\]

where \( \Delta b_{\perp} \) is the average interval of orthogonal baselines.

**III. THREE-DIMENSIONAL SAR FOCUSING VIA COMPRESSIVE SENSING**

To reconstruct the reflectivity function \( \gamma(s) \) from \( g = [g_1 \ g_2 \ \ldots \ g_M]^T \) is a semidiscrete problem. In practical applications, the formulation of SAR tomography can be modeled as follows by discretizing \( \gamma(s) \) [3], [5]:

\[
g = A \cdot \gamma.
\]

The row of \( A \) is defined as

\[
a_m = [\exp(j2\pi \xi_m s_1) \ \exp(j2\pi \xi_m s_2) \ \ldots \ \exp(j2\pi \xi_m s_N)]
\]

is

\[
\gamma = [\gamma(s_1) \ \gamma(s_2) \ \ldots \ \gamma(s_N)]^T
\]

\( \gamma(s_n) \) is the samples of \( \gamma(s) \) at nsr position \( s_n \), \( N \) is the sampling number of nsr. The sampling interval of nsr is \( \Delta s = \eta \rho_s \), \( \eta \) is a number larger than one [3], [4]. Because \( M \), the available number of passes achieved by repeat-pass missions of single-antenna SAR systems, is smaller than \( N \), (6) expresses an underdetermined system.

In one azimuth-range resolution cell of SAR image, particularly in urban environment, the overall reflectivity is usually dominated by only a few of dominating scatterers resulting from metallic structures or specular and dihedral reflections [5]. Hence, \( \gamma \) is sparse in the identity orthogonal basis. Considering the sparsity of \( \gamma \), [3]–[5] proposed a novel SAR tomographic focusing method based on CS, which is a new and popular approach for sparse signal reconstruction from significantly fewer measurements than that were traditionally thought necessary [6].

If we take the \( \ell_0 \) norm \( ||\gamma||_0 \) as the measure of sparsity, CS recovers \( \gamma \) by solving [13]

\[
\hat{\gamma} = \arg \min_{\gamma} ||\gamma||_0 \quad \text{subject to} \quad g = A \cdot \gamma.
\]

Because (9) is intractable in general, some methods have been proposed. For examples, method of frames (MOF) [14], best orthogonal basis (BOB) [15], matching pursuit (MP) [16], basis pursuit (BP) [13], and so on. BP finds signal representations by convex optimization. This optimization principle leads to decompositions that have very different properties from the MOF, in particular, they can be much sparser. Because BP is based on global optimization, it is stably super-resolve in ways that MP cannot. Because BP always delivers the decomposition in an optimal basis, it seems better than the BOB method in resolving nonorthogonal structures. According to these, BP method is chosen to solve the problem of SAR tomography. BP minimizes the \( \ell_1 \) norm instead of \( \ell_0 \) norm.

\[
\hat{\gamma} = \arg \min_{\gamma} ||\gamma||_1 \quad \text{subject to} \quad g = A \cdot \gamma.
\]

In order to prove whether CS is applicable for SAR tomography or not, we should check the property of matrix \( A \) in addition to the sparsity of \( \gamma \). Now, the traditional way is by means of the sufficient condition of restricted isometry property (RIP) [7]. However, it is intractable to prove whether a given matrix satisfies RIP. It has been proved that \( A \) is in some sense ideal for recovering sparse signal if it is a certain random matrix. For example, the properly normalized Gaussian matrix with i.i.d. entries, the matrix obtained by randomly sampling rows of DFT matrix [8], [9]. More generally, a matrix obtained by randomly sampling rows of an orthonormal matrix will be a good sensing matrix [10], [11].

In SAR tomography, \( \xi_m \), which is determined by \( b_{\perp m} \), is usually located uniformly at random in certain continuous frequency band, instead of on an equipased lattices. As a result, the measurements are random samples of continuous frequency spectrum of \( \gamma(s) \), rather than random samples of DFT of \( \gamma(s) \). \( A \) does not consist of randomly sampling rows of an orthonormal matrix. Hence, it is not sufficient to prove the applicability of CS for SAR tomography using the theory in [10] and [11]. Recently, [12] introduced a simple and very general theory of compressive sensing. The main contribution of this theory is to provide a simple framework which applies to not only all the standard compressive sensing models but also
some new ones, including the random samples in continuous frequency spectrum. In this theory, the sensing mechanism simply selects sensing vectors $\mathbf{a}$ independently at random from a probability distribution $F$. Candes proved that if the probability distribution $F$ obeys a simple incoherence property and an isotropy property, we can faithfully recover sparse signals from a minimal number of measurements. Thus, we use this theory to prove the applicability of CS for SAR tomography.

- **Isotropy property**: We say that $F$ obeys the isotropy property if

$$E(\mathbf{a}^T \mathbf{a}^*) = I. \quad (11)$$

In other words, the isotropy condition states that the components of $\mathbf{a} \sim F$ have unit variance and are uncorrelated.

- **Incoherence property**: We may take the coherence parameter $\mu(F)$ to be the smallest number satisfying the following equation:

$$\max_{1 \leq n \leq N} |a[n]|^2 \leq \mu(F) \quad (12)$$

where $a[n]$ is the $n$th entry of $\mathbf{a}$. The smaller $\mu(F)$, i.e., the more incoherent the sensing vectors, the fewer samples we need for accurate recovery.

- **Theorem**: If $F$ obeys isotropy or “near isotropy,” i.e., $E(\mathbf{a}^T \mathbf{a}^*) \approx I$, and pick any scalar $\beta$. Then, with probability at least $1 - 5/N - e^{-\beta}$, $\gamma$ is the unique minimize to $\gamma$ provided that

$$M \geq C_{\beta} \cdot \mu(F) \cdot K \cdot \log N \quad (13)$$

where $K$ is the sparsity of $\gamma$. More precisely, $C_{\beta}$ may be chosen as $C_0(1 + \beta)$ for some positive numerical constant $C_0$.

In the repeat-pass missions of SAR system, the orthogonal baseline is usually located uniformly at random over a certain overall orthogonal baseline length. For example, the orthogonal baselines of Envisat used in section 4 of this letter are located uniformly at random over $[−706.698]$. The orbit of TerraSAR-X is controlled in a predefined tube of 500 m diameter throughout the entire mission [5]. Hence, $\xi$ follows uniform distribution; the probability density function is

$$f(\xi) = \lambda r/2B_\perp. \quad (14)$$

Then a simple calculation we shall not detail shows that the entry of $E(\mathbf{a}^T \mathbf{a}^*)$ is

$$\begin{align*}
[E(\mathbf{a}^T \mathbf{a}^*)]_{xy} &= E \{\exp(j2\pi \xi(s_x - s_y))\} \\
&= \int \exp[j2\pi \xi(x - y)\Delta s] \cdot f(\xi) \cdot d\xi \\
&= \begin{cases}1, & x = y \\
\sin c(\pi(x - y)/\eta), & x, y \in [1, N], \ x \neq y.
\end{cases}
\end{align*} \quad (15)$$

The maximum is about 0.2 when $x \neq y$, so $E(\mathbf{a}^T \mathbf{a}^*) \approx I$.

In other words, this distribution is near isotropic. Because

<table>
<thead>
<tr>
<th>Wave</th>
<th>Distance from scene center</th>
<th>Azimuth resolution</th>
<th>Range resolution</th>
<th>Incidence angle</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.056m</td>
<td>843130m</td>
<td>4m</td>
<td>8 m</td>
<td>21</td>
</tr>
</tbody>
</table>

Fig. 2. Distribution of orthogonal baselines.

Fig. 3. Envisat-ASAR image of Angel Stadium.

$\mathbf{a}[n] = \exp(j2\pi \xi s_n)$, the incoherence $\mu(F)$ is 1. According to the aforementioned theorem, we can perfectly recover $\gamma$ when $M$ is on the order of $K \cdot \log N$.

The applicability of CS for SAR tomography has been proved theoretically using the latest improvement of CS in [12]. In the next section, we will apply CS-based focusing method to ENVISAT-ASAR data to confirm its applicability farther.

### IV. Experimental Results

The CS-based focusing method of SAR tomography is applied to the 3-D reconstruction of Angel Stadium of Anaheim (N 33°48′0.11″, W 117°52′58.85″), California, America. The data set consists of 20 images acquired by Envisat-ASAR in the period from October 29th 2003 to November 7th 2007. The image of June 15th 2005 is selected as the reference image. The main system parameters are presented in Table I. Fig. 2 shows the distribution of orthogonal baselines. It is obvious that some large spatial gaps are presented for orthogonal baselines. The overall orthogonal baseline length $B_\perp$ is about 1403 m, leading to a Rayleigh resolution $\rho_x = 16.8$ m. The average interval of orthogonal baseline $\Delta b_\perp$ is about 79 m. According to (5), the

<table>
<thead>
<tr>
<th>Table I</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Parameters</td>
</tr>
<tr>
<td>Wave length</td>
</tr>
<tr>
<td>0.056m</td>
</tr>
</tbody>
</table>
critical extent of scene $2s_{\text{max}} = 320$ m. Fig. 3 gives the SAR image of Angel Stadium. In our processing, the phase errors associated to atmospheric variations and ground deformations are compensated using the method proposed in [17].

Because we do not know the exact elevation of the ground in this area, a pixel marked by the star in Fig. 3 is selected as the reference point. Fig. 4 shows the focusing result of this pixel, the result has been converted to the height perpendicular to ground according to the incidence angle. This pixel is not near the stadium, so there is only one dominating scatterer near ground inside this pixel. The height of this dominating scatterer is set to be zero, and all the following results are corrected with reference to it. We carry out the tomographic processing of three azimuth-height sections that are shown by the dashed lines in Fig. 3.

First, to show what we expect from the tomographic processing of the three sections, we forecast the focusing results according to the intersections between the three sections and the 3-D model of the stadium, this method is firstly adopted in [3], [4]. The forecasted results are shown in Fig. 5. Section 1 and section 2 cross the wall and coverage of the stadium, most of the dominating scatterers should be located at the positions shown by the dashed lines in Fig. 5(a) and (b). Section 3 crosses the bleachers. Because the seats resemble dihedral reflections, the dominating scatterers should be presented in the rectangle lined out by the dashed line in Fig. 5(c).

Fig. 6 presents the focusing results of the three sections using SVD. Fig. 7 presents the focusing results of the three sections using CS. In processing of CS, the sampling interval of nsr $\Delta s$ is 3.2 m, and the sampling number of nsr $N$ is 100. According to the discussion in section 3, the twenty images are sufficient for good recovery of four scatterers. In Fig. 7, the focusing results using CS are in good agreement with the forecasted ones.
Fig. 7. Focusing results of CS-based focusing method. (a) Section 1. (b) Section 2. (c) Section 3.

Fig. 8. Height of Angel Stadium resulting from the focusing results in Fig. 7.

V. Conclusion

In SAR tomography, the measurements are random samples of continuous frequency spectrum, instead of random samples of DFT. We proved the applicability of CS for SAR tomography theoretically using the latest improvement of CS, which provides a simple framework for the measurement strategy in SAR tomography. Satisfying results were achieved in the processing of Envisat-ASAR data over Angel Stadium of Anaheim. Both the theoretical analysis and satisfying results of real data processing validated the CS-based SAR tomographic focusing method proposed in [3]–[5].

ACKNOWLEDGMENT

The authors thank J. Jasper from America for making the 3-D model of Angel Stadium.

REFERENCES