Enhanced ISAR imaging by exploiting the continuity of the target scene

Zhao, Lifan; Wang, Lu; Bi, Guoan; Wan, Chunru; Yang, Lei

2013

Wang, L., Zhao, L., Bi, G., Wan, C., & Yang, L. (2013). Enhanced ISAR Imaging by Exploiting the Continuity of the Target Scene. *IEEE Transactions on Geoscience and Remote Sensing*, 52(9), 5736 - 5750.

<https://hdl.handle.net/10356/104851>

<https://doi.org/10.1109/TGRS.2013.2292074>

© Copyright 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. The published version is available at: [<http://dx.doi.org/10.1109/TGRS.2013.2292074>].

Downloaded on 26 Apr 2023 21:22:30 SGT

Enhanced ISAR Imaging by Exploiting the Continuity of the Target Scene

Lu Wang, Lifan Zhao, Guoan Bi, *IEEE Senior Member*, Chunru Wan and Lei Yang

Abstract—This paper presents a novel inverse synthetic aperture radar (ISAR) imaging method by exploiting the inherent continuity of the scatterers on the target scene to obtain enhanced target images within a Bayesian framework. A simplified radar system is utilized by transmitting the sparse probing frequency signal, where the ISAR imaging problem can be converted into one dealing with under-determined linear inverse scattering. Following the Bayesian compressive sensing theory, a hierarchical Bayesian prior is employed to model the scatterers in the range-Doppler plane. In contrast to the independent prior on each scatterer in conventional Bayesian compressive sensing, a correlated prior is proposed to statistically encourage the continuity structure of the scatterers in the target region. To overcome the intractability of the posterior distribution, the Gibbs sampling strategy is used for Bayesian inference. The parameters of the signal model are inferred efficiently from samples obtained by the Gibbs sampler. Because the proposed method is a data-driven learning process, the tedious parameter tuning process required by the convex optimization based approaches can be avoided. Both the synthetic and the experimental results demonstrate that the proposed algorithm can achieve substantial improvements in the scenarios of limited measurements and low signal-to-noise ratio compared with other reported algorithms for ISAR imaging problems.

Index Terms—ISAR imaging, Bayesian compressive sensing, Gibbs sampler, structure of the continuity.

I. INTRODUCTION

HIGH resolution radar imaging based on sparse representation has recently received much attention. It has been shown in [1]–[8] that high-resolution image can be achieved with a limited number of pulses by reformulating the ISAR or SAR imaging problem into an under-determined linear inverse scattering problem. Taking the uncertainties of the under-determined inverse problem into consideration, three main strategies have been proposed to solve the problem under the sparsity constraint of the solution. One outstanding class of these methods is to reconstruct the sparse signal by using optimization strategies. Such methods are considered to solve an l_p ($0 \leq p \leq 1$) minimization problem, known as competition-optimization techniques, such as the FOCal Under-determined System Solver (FOCUSS) [9], the Basis Pursuit (BP) [10], and the Least Absolute Shrinkage and Selection Operator (LASSO) [11]. Another class of methods to achieve sparse solutions is known as iterative greedy algorithms, such as Matching Pursuit (MP) [12] and its variations [13]–[15]. Sparse

recovery problem can be also reformulated within a Bayesian framework, where the estimation of the sparse signal is given by the mean of the posterior distribution. The efficient inference of the model parameters can be performed by the iterative learning methods such as Expectation Maximization (EM) algorithm [18], variational Bayesian (VB) inference [19] [20] and stochastic sampling methods [21]. Sparse recovery techniques under the Bayesian framework are referred as Bayesian compressive sensing (BCS) in the literature.

All these above-mentioned methods have been used in radar imagery applications via sparse representation. By exploiting the sparsity of ISAR/SAR images, sparsity based imagery methods are shown to outperform the conventional range-Doppler algorithm (RDA) [16] in terms of achieving high resolution with fewer measurements. The l_1 -norm based BP is used to obtain a high-resolution image with fewer pulses in ISAR [2][8] and in SAR [23]. Apart from the l_1 -norm regularization, additional constraints and procedures are added for other purposes. For example, the Total Variation (TV) constraint is added along with the l_1 -norm constraint [1][22] to deal with the speckle noise, and an iterative procedure for estimating the phase error is incorporated [24] to obtain a focused image. The MP is used for imagery of the target with uniform or even maneuvering rotation [3] due to its computational efficiency and the BCS strategy has been found to obtain an auto-focused ISAR image [4][5] and high-resolution SAR image [6]. To impose the sparsity, the Laplace prior is employed for each scatterer on the target scene where the posterior distribution is intractable [4] and an estimate of the sparse signal is obtained by maximizing its posterior probability via a quasi-Newton technique. In [6], a hierarchical prior which facilitates a tractable posterior distribution is used and the sparse signal is inferred by the EM algorithm. In the recent literature of ISAR imagery based on sparse representation, undersampling is performed for each range cell in the cross-range domain after the pulse compression. Instead of transmitting a wide-band signal and performing the conventional pulse compression in the range domain, a different and simplified radar system which transmits the sparse probing frequencies (SPF) is proposed to obtain a high-resolution ISAR image by using the convex optimization technique with l_1 -norm constraint [8].

In general, ISAR images of the practical targets always exhibit strong continuity. More concretely, nonzero scatterers on the target region are continuously located along either the range or the cross-range domain in the target scene. Considering that radar transmits sparse probing frequencies as used in [8], we reformulate the ISAR imaging problem

The corresponding author is Dr Guoan Bi. Lu Wang, Lifan Zhao, Guoan Bi and Lei Yang are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, e-mail: wa0001lu, zhao0145, egbi and yanglei@ntu.edu.sg. Chunru Wan is with DSO, Singapore, e-mail: wchunru@dso.org.sg

into the Bayesian framework to exploit the inherent continuity of the scatterers in the target scene for enhanced ISAR images. In contrast to the work reported in [4] and [6], where independent prior for each scatterer in target scene is simply used to impose the sparsity of the target image, we propose a correlated prior to enforce the continuity of the image. A similar idea has been explored for recovering a tree-structured sparse signal by compressive sensing [21], cluster-structured sparse signals in [26] and the ultra sound image by using the Markov random field (MRF) to exploit the spatial correlation of the image pixels in the Fourier domain [27]. These methods exploiting the inherent structure of the signal can be generalized as model-based compressive sensing [28]. Some research focusing on different structures of the signal has been reported, such as block-sparsity [29]-[32] where a rigid signal structure is directly imposed on the signal vector itself, tree-structure [20][21][33][34] and other spatial correlation [26][27] where structures are statistically imposed on the Bayesian prior to offer a flexible modeling of the signal structure. Although the prior knowledge about the structure of the signal and its specific imposing strategy can be quite different, the fundamental principle to exploit the structure of the signal to achieve the guarantees of better performance remains the same. Inspired by the similar idea, this paper introduces the continuity structure of the target images into the ISAR imagery problem within the Bayesian framework. By exploiting the continuity structure of the target image, the proposed imagery algorithm offers the advantages of high recovery accuracy and resolution in low signal-to-noise ratio (SNR) environments with fewer measurements.

This paper is organized as follows. Section II briefly reviews the principles of the ISAR imaging and the mathematical signal model based on sparse probing frequencies. Section III introduces continuity patterns of the scatterers on the target scene and constitutes the Bayesian model to enforce such continuities. The efficient inference of the parameters and hidden variables via the Markov chain Monte Carlo (MCMC) sampling is given in Section IV. Results of experiments with synthetic and practical data are presented in Section V to show the effectiveness of the proposed algorithm. Conclusions are drawn in Section VI.

Throughout the paper, we denote the variables by lowercase and uppercase letters, e.g., k and K , the vectors by bold lowercase letters, e.g., $\mathbf{u}$ and its l th element $\mathbf{u}(l)$, and the matrices by bold uppercase letters, e.g., $\mathbf{D}$ and its l th column by $\mathbf{d}_l$. Denote $\|\bullet\|_q$ for $q = 0, 1, 2$ the l_0 , l_1 and l_2 -norm. $\mathbf{u} \in \mathbb{C}^{n \times 1}$ represents a column vector of length n with complex elements.

II. ISAR IMAGING

A. Signal model for ISAR imaging

The ISAR imaging geometry for a plane target is shown in Fig. 1 [17], where three coordinate systems T_z , T_x and T_y are used. T_z represents the reference system embedded in the radar with the axis z_2 oriented along the line of sight (LOS). T_x is a time varying coordinate system chosen to keep the x_2 axis oriented along the LOS, the x_3 axis along the effective

Fig. 1. The geometry of the radar imaging of a target when $t = 0$.

rotation vector Ω_{eff} and the origin fixed at the target center O . T_y is a reference system embedded in the target which is coincident with T_x when and only when $t = 0$. The distance between target and radar is R_0 . Ω_T is the total angular rotation vector caused by the translation and angular motions of the target, and Ω_{eff} is the effective rotation vector which is the projection of Ω_T on the plane orthogonal to the LOS. Both Ω_T and Ω_{eff} are applied to the target center O . P is the k th scatterer with position vector $\mathbf{y}_k = [y_{1,k} \ y_{2,k} \ y_{3,k}]^T$ in T_y coordinate system. The Doppler induced by the rotation of scatterer P is generated by [17]

$$x_2(\mathbf{y}_k, t) = a_2(\mathbf{y}_k) + b_2(\mathbf{y}_k) \cos(|\Omega_T|t) + \frac{c_2(\mathbf{y}_k)}{|\Omega_T|} \sin(|\Omega_T|t) \quad (1)$$

where

$$\begin{aligned} a_2(\mathbf{y}_k) &= \frac{(\Omega_T \cdot \mathbf{y}_k)}{|\Omega_T|^2} \Omega_{T_2}, \\ b_2(\mathbf{y}_k) &= y_{2,k} - \frac{(\Omega_T \cdot \mathbf{y}_k)}{|\Omega_T|^2} \Omega_{T_2}, \\ c_2(\mathbf{y}_k) &= |\Omega_{eff}| y_{1,k}, \end{aligned}$$

$|\Omega|$ is the modulus of vector Ω , and Ω_{T_2} is the coordinate of Ω_T along the x_2 axis. For target with a small constant angular velocity, (1) can be approximated by its first-order polynomial:

$$x_2(\mathbf{y}_k, t) = y_{2,k} + |\Omega_{eff}| y_{1,k} t \quad (2)$$

and the cross-range profile in a particular range cell can be rewritten as

$$\begin{aligned} s(\tau, t) &= \sum_{k=1}^K a_k B \cdot \text{rect}\left(\frac{\tau}{T_{ob}}\right) \\ &\times \text{sinc}\left[B\left(\tau - \frac{2}{c}(x_2(\mathbf{y}_k, t) + R_0)\right)\right] e^{j2\pi(\gamma_k t)} \end{aligned} \quad (3)$$

where

$$\gamma_k = \frac{2F_c}{c}(x_2(\mathbf{y}_k, t) + R_0),$$

a_k is the complex amplitude of the scatterer k , B is the transmitted signal bandwidth, F_c is the carrier frequency, c is the speed of light, T_{ob} is the observation time and K is the number of point scatterers in a particular range cell, **rect** and **sinc** denote the rectangular and the sinc functions, respectively. Expression $\frac{2|\Omega_{eff}|y_{1,k}}{\lambda}$ in (3) denotes the Doppler frequency of scatterer k , where $\lambda = c/F_c$. The conventional RDA has been used to obtain the target image based on the signal model described in (3).

B. Sparse probing frequencies

With the compressive sensing theory, it is unnecessary for Radar to transmit a wide-band signal because a few probing frequencies are sufficient to reconstruct the ISAR image by solving a convex optimization problem with l_1 -norm constraint as demonstrated in [8]. Therefore, a simplified Radar system which transmits the sparse probing frequency (SPF) signal is proposed to obtain a high resolution ISAR image. It is shown in [8] that if a probing frequency signal $\exp(j2\pi f_l \tau)$ with frequency f_l at time t_l is transmitted, the received signal at the range frequency $f_l \in [F_c - B/2, F_c + B/2]$ and cross-range time $t_l \in [1, L]$ is given by

$$\mathbf{r}(l) = \sum_{k=1}^K a_k \exp \left(-j4\pi f_l \frac{R_0 + y_{2,k} + |\Omega_{eff}| y_{1,k} t_l}{c} \right). \quad (4)$$

Discretizing the target scene of interest into an $M \times N$ grid in the cross-range and range domains as illustrated in Fig. 2, each scatterer located at grid (m, n) is denoted by $(y_{2,m}, y_{1,n})$ with amplitude $a_{m,n}$, where $y_{1,n}$ in the cross-range domain belongs to

$$\left(-\frac{ND_r}{2} : D_r : \frac{ND_r}{2} - D_r \right) \quad (5)$$

with $D_r = \frac{F_D}{N} \frac{\lambda}{2|\Omega_{eff}|}$, F_D is the maximal unambiguous Doppler shift which is determined by the pulse repetition frequency (PRF), and $y_{2,m}$ in the range domain belongs to $\left(-\frac{MR_r}{2} : R_r : -\frac{(M-1)R_r}{2} \right)$ with $R_r = c/2B$. The resolution is defined by (R_r, D_r) . An over-completed dictionary $\mathbf{D}$ of

Fig. 2. Cross-range and range domain with a designed resolution.

size $L \times MN$ can be constructed according to the target scene in Fig. 2 with atom $\mathbf{d}_{m,n}$ whose element is defined as

$$\mathbf{d}_{m,n}(l) = \exp \left\{ -j4\pi \frac{f_l}{c} (R_0 + y_{2,m} + |\Omega_{eff}| y_{1,n} t_l) \right\}. \quad (6)$$

Denote the coefficient vector

$$\mathbf{a}^T = [a_{1,1} \dots a_{m,1} \ a_{1,2} \dots a_{1,n} \dots a_{m,n}],$$

we have

$$\mathbf{r} = \mathbf{D}\mathbf{a}. \quad (7)$$

Because $L < MN$, (7) is under-determined. However, since vector $\mathbf{a}$ is sparse due to the fact $K \ll MN$, there exists a unique solution to (7) that can be obtained by sparse recovery methods provided that the dictionary $\mathbf{D}$ satisfies the restricted isometry property (RIP)[8], [38].

C. Maximal mutual coherence of $\mathbf{D}$

Since the RIP property of a dictionary is generally hard to analyze, instead we analyze theoretically the maximal mutual coherence of the dictionary in the Appendix when full measurements are used. As shown in the Appendix, the maximal coherence of the dictionary varies with ratio B/F_c for a given N . For X-band signal which is commonly used in ISAR, we generally have $0.04 \leq \frac{B}{F_c} \leq 0.2$. By designing the radar system with a proper B/F_c , a dictionary with relatively small maximal coherence can be obtained.

However, random sampling in the range domain will result in a dictionary with increased maximal mutual coherence. The maximal coherence of $\mathbf{D}$ depends on both the number of samples used and their positions. Therefore, it is hard to theoretically analyze the maximal mutual coherence under random sampling. We evaluate the maximal mutual coherence of the dictionary $\mathbf{D}$ versus the number of random samples by Monte Carlo method for fixed B/F_c and N in section V-A. As shown later, the maximal mutual coherence of $\mathbf{D}$ is either similar to or smaller than that of the random Gaussian dictionary of the same size which has been found to satisfy the RIP with a high probability [38].

III. CONTINUITY PATTERNS AND BAYESIAN MODEL

The conventional CS-based methods have used only the sparse property of ISAR and SAR images to solve the radar imagery problems [1]-[8]. In addition to the sparsity, we further exploit the inherent continuity of the target scatterers in either the range or the cross-range domain to obtain an enhanced ISAR image under the Bayesian framework. Following the work reported in [8], let us consider a simple radar system which transmits the sparse probing frequencies instead of the wide-band linear frequency modulated (LFM) signal. Before reformulating the ISAR imagery problem in the Bayesian framework, let us first introduce three continuity patterns of the scatterers in range and cross-range domains that are to be exploited in this paper.

A. Continuity Patterns

In real world scenarios, the target region is continuous in the range and cross-range domains. It is therefore assumed in this paper that the scatterers of the target are not isolated, but located continuously in the target scene. This assumption is reasonable for ISAR image since the target of interest in ISAR is usually small and can be regarded as a mass of small scatterers. This phenomenon can also be observed in the practical ISAR data.

The continuity patterns for a particular scatterer indexed by (m, n) in the target scene are given in Fig. 3, where the shaded squares denote the scatterers with nonzero amplitudes and the white ones are the scatterers which are irrelevant to the scatterer (m, n) . The amplitudes of the white ones can be nonzero or zero. If any of the patterns in Fig. 3 (a), (b) and (c) is observed, it is reasonable to assume that $a_{m,n}$ for scatterer (m, n) is nonzero with a high probability. Therefore the continuity either in range domain, cross-range domain, or in both domains can be encouraged accordingly.

Fig. 3. Three kinds of continuity patterns

B. Bayesian Model

Under the framework of Bayesian compressive sensing, a hierarchical prior is used to enforce the sparsity in the signal coefficient vector $\mathbf{a}$, where elements in $\mathbf{a}$ are independent and each element $a_{m,n}$ follows a Bernoulli-Gaussian distribution, which is also known as a “spike-and-slab” prior [20][25][35]:

$$p(a_{m,n} | q_{m,n}, \beta_{m,n}) = \sum_{\varepsilon \in \{0,1\}} p(a_{m,n} | w_{m,n} = \varepsilon, \beta_{m,n}) Pr(w_{m,n} = \varepsilon | q_{m,n}), \quad (8)$$

where the hyperparameter $q_{m,n}$ is the prior probability of having a non-zero $w_{m,n}$ which is a hidden binary variable indicating if $a_{m,n}$ is nonzero or not and follows the Bernoulli distributions with parameter $q_{m,n}$

$$Pr(w_{m,n} | q_{m,n} = 1) = q_{m,n},$$

and

$$Pr(w_{m,n} | q_{m,n} = 0) = 1 - q_{m,n}.$$

Also $p(a_{m,n} | w_{m,n} = 1, \beta_{m,n})$ follows $\mathcal{CN}(a_{m,n} | 0, \beta_{m,n}^{-1})$ when $w_{m,n} = 1$, where $\mathcal{CN}(0, \beta_{m,n}^{-1})$ denotes the complex Gaussian distribution with zero mean and precision $\beta_{m,n}$. The conditional prior of $a_{m,n}$ can be directly obtained by

$$p(a_{m,n} | q_{m,n}, \beta_{m,n}) = (1 - q_{m,n}) \delta(|a_{m,n}|) + \frac{q_{m,n} \beta_{m,n}}{\pi} \exp(-\beta_{m,n} |a_{m,n}|^2). \quad (9)$$

With the prior in (9) on each element of the signal vector $\mathbf{a}$, the distribution of $\mathbf{a}$ given $\mathbf{q}^T = [q_{1,1} \cdots q_{M,N}]$ and $\beta^T = [\beta_{1,1} \cdots \beta_{M,N}]$ is

$$p(\mathbf{a} | \mathbf{q}, \beta) = \prod_{i=1}^{MN} \left[(1 - q_i) \delta(|a_i|) + \frac{q_i \beta_i}{\pi} \exp(-\beta_i |a_i|^2) \right] \quad (10)$$

where $1 \leq i \leq MN^1$.

Despite the sparsity, the scatterer distribution of a practical target image also exhibits a strong continuity which has not been exploited in most CS-based ISAR and SAR imagery such as the method proposed in [6]. In our framework, the correlated priors for $\beta_{m,n}$ and $q_{m,n}$ are proposed to encourage the continuity, which is similar to the model used in [20] and [26], where the structural information is statistically imposed on the Bayesian prior rather than explicitly making a hard imposition of the structure on the signal vector [28][29]. In this way, signals are not restricted by the imposed structure and the corresponding algorithms would be more flexible and robust for signals with weak structural information.

1) *Correlated prior for β* : Let us assume $a_{m-1,n}$, $a_{m+1,n}$, $a_{m,n-1}$ and $a_{m,n+1}$ are all nonzero. To enforce continuity pattern in both range and cross-range domains in Fig. 3 (c), it is proper to choose a correlated prior for the precision of the scatterers in sets $\{j = n | m-1 \leq i \leq m+1\}$ and $\{i = m | n-1 \leq j \leq n+1\}$, i.e., they share the same precision:

$$\beta_{m,n} = \beta_{m-1,n} = \beta_{m+1,n} = \beta_{m,n-1} = \beta_{m,n+1}.$$

To obtain a tractable posterior distribution of $\beta_{m,n}$, we further impose a Gamma prior, $\text{Gamma}(b, d)$, on $\beta_{m,n}$ with parameters b and d . Since the Gamma prior is conjugate to Gaussian distribution with an unknown precision, the posterior distribution of $\beta_{m,n}$ given $\mathbf{a}$ also follows a Gamma distribution²:

$$\begin{aligned} p(\beta_{m,n} | \mathbf{a}, b, d) &= p(\beta_{m,n} | \bar{\mathbf{a}}_{m,n}, b, d) \\ &= \text{Gamma}\left(b + 5, d + \|\bar{\mathbf{a}}_{m,n}\|_2^2\right) \end{aligned} \quad (11)$$

with $\bar{\mathbf{a}}_{m,n}$ defined as the neighbor set of $a_{m,n}$:

$$\bar{\mathbf{a}}_{m,n}^T = [a_{m,n} \quad a_{m-1,n} \quad a_{m+1,n} \quad a_{m,n-1} \quad a_{m,n+1}].$$

Similarly, the posterior distribution of $\beta_{m,n}$ for continuity pattern (a) in range domain and continuity pattern (b) in cross-range domain can be derived as follows:

$$p(\beta_{m,n} | \mathbf{a}, b, d) = \text{Gamma}\left(b + 3, d + \|\bar{\mathbf{a}}_{m,n}^k\|_2^2\right) \quad (12)$$

with $k = 1, 2$ and $\bar{\mathbf{a}}_{m,n}^1 = [a_{m,n} \quad a_{m-1,n} \quad a_{m+1,n}]^T$ for continuity pattern (a) and $\bar{\mathbf{a}}_{m,n}^2 = [a_{m,n} \quad a_{m,n-1} \quad a_{m,n+1}]^T$ for continuity pattern (b).

If none of the continuity patterns occurs, then $\beta_{m,n}$ is assumed to be independent to others. By exerting a Gamma prior to it, the posterior distribution of $\beta_{m,n}$ can be easily calculated by

$$p(\beta_{m,n} | a_{m,n}, b, d) = \text{Gamma}(b + 1, d + |a_{m,n}|^2). \quad (13)$$

2) *Prior for hidden variable $\mathbf{q}$* : To incorporate the continuity information in the inference of the hidden variable $\mathbf{q}$, three kinds of sparsity patterns, similar to those in [21] and [26], are introduced. They are the patterns for strong acceptance, strong rejection and weak rejection, respectively. To facilitate the inference, a Beta prior, which is conjugate to Bernoulli distribution with three different sets of parameters $\{e_k, f_k\}_{k=0,1,2}$, is assigned for $w_{m,n}$ according to the introduced sparsity patterns. The sparsity patterns are defined below based on the idea of continuity:

Strong rejection (referred as Sparsity Pattern 0 (SP0)): If $\bar{\mathbf{a}}_{m,n} \setminus a_{m,n}$, i.e., the sub-vector of $\bar{\mathbf{a}}_{m,n}$ excluding the element of $a_{m,n}$, are all zeros, it is highly possible that $a_{m,n}$ is also zero due to the continuity of the target scene. The prior $\text{Beta}(e_0, f_0)$ with $e_0 < f_0$ is used to reject scatterer (m, n) and to guarantee the probability, $q_{m,n}$, for $w_{m,n} = 1$, being small.

²Gamma distribution: $f(x; b, d) = d^b x^{b-1} e^{-dx} / \Gamma(b)$, where $\Gamma(b)$ is the Gamma function evaluated at b .

¹We abuse the subscripts i and m, n for easy representation.

Strong acceptance (referred as Sparsity Pattern 2 (SP2)): If any of the continuity patterns in Fig. (3) is observed, then it is highly possible that $a_{m,n}$ is nonzero to encourage the continuity of the target image. The prior $Beta(e_2, f_2)$ with $e_2 > f_0$ is used to accept scatterer (m, n) and to guarantee the probability, $q_{m,n}$, for $w_{m,n} = 1$, being large.

Weak rejection (referred as Sparsity Pattern 1 (SP1)): Apart from those situations mentioned in SP0 and SP2, no apparent conclusion can be drawn on the value of $q_{m,n}$. The prior $Beta(e_1, f_1)$ with $e_1 = f_1$ is used to exert non-informative prior on the $q_{m,n}$.

Since Beta distribution is conjugate to Bernoulli for its unknown parameter $q_{m,n}$, the posterior distribution for $q_{m,n}$ also follows a Beta distribution given by:

$$p(q_{m,n}^k | \mathbf{a}, e_k, f_k) = p(q_{m,n}^k | \bar{\mathbf{a}}_{m,n}, e_k, f_k) = \mathbf{Beta}(e_k + \|\bar{\mathbf{a}}_{m,n}\|_0, f_k + 5 - \|\bar{\mathbf{a}}_{m,n}\|_0) \quad (14)$$

where $k \in \{0, 1, 2\}$ and $\|\mathbf{a}\|_0$ gives the number of the nonzero components in $\mathbf{a}$.

IV. PROPOSED IMAGERY ALGORITHM

A. Sparse Solution and the Gibbs Sampler

Suppose that the signal is corrupted by a complex white Gaussian noise $\mathbf{n}$ with zero mean and unknown precision β_0 , then the measurement vector $\mathbf{r}$ is

$$\mathbf{r} = \mathbf{D}\mathbf{a} + \mathbf{n} \quad (15)$$

where $\mathbf{n} \in \mathbf{C}^{L \times 1}$ and follows $\mathcal{CN}(0, \beta_0^{-1} \mathbf{I})$. The posterior distribution of all the variables given $\mathbf{r}$,

$$p(\mathbf{a}, \mathbf{q}, \beta, \beta_0 | \mathbf{r}) \propto p(\mathbf{r} | \mathbf{D}\mathbf{a}, \beta_0) p(\beta_0) p(\mathbf{a} | \mathbf{q}, \beta) p(\mathbf{q}) p(\beta), \quad (16)$$

is intractable due to the implicit priors for β and $\mathbf{q}$ and no analytical estimation can be found for $\{\mathbf{a}, \mathbf{q}, \beta, \beta_0\}$. The Gibbs sampler [36], the standard MCMC sampling, is used to generate samples asymptotically distributed according to (16) and to find the estimation of $\{\mathbf{a}, \mathbf{q}, \beta, \beta_0\}$. The sequential sampling procedures of the Gibbs sampler are given in the following four steps:

- Sampling a_i according to its posterior distribution;
- Sampling β_i according to its posterior distribution given in (11), (12) and (13) for different continuity patterns;
- Sampling q_i according to its posterior distribution given in (14) for different sparsity patterns;
- Sampling noise precision β_0 according to its posterior distribution.

In step (a), the posterior distribution of a_i is required to be derived. Based on (10) and (15), the posterior distribution of $\mathbf{a}$ is proportional to

$$p(\mathbf{a} | \mathbf{r}, \mathbf{q}, \beta, \beta_0) \propto p(\mathbf{r} | \mathbf{D}\mathbf{a}, \beta_0) \times \left\{ \prod_i^{MN} \left[(1 - q_i) \delta(|a_i|) + \frac{q_i \beta_i}{\pi} \exp(-\beta_i |a_i|^2) \right] \right\}. \quad (17)$$

By (17), one can easily compute the posterior distribution $p(a_i)$ of a_i given $\{\mathbf{a} \setminus a_i, \mathbf{r}, \mathbf{q}, \beta, \beta_0\}$ for $1 \leq i \leq MN$, which follows

$$p(a_i | \mathbf{a} \setminus a_i, \mathbf{r}, \mathbf{q}, \beta, \beta_0) \propto (1 - \tilde{q}_i) \delta(|a_i|) + \tilde{q}_i \mathcal{CN}(a_i | \tilde{\mu}_i, \tilde{\beta}_i^{-1}) \quad (18)$$

where

$$\begin{aligned} \tilde{\beta}_i &= \beta_i + \beta_0 \mathbf{d}_i^H \mathbf{d}_i, \\ \tilde{\mu}_i &= \tilde{\beta}_i^{-1} \beta_0 \mathbf{d}_i^H \left(\mathbf{r} - \sum_{j=1, j \neq i}^{MN} a_j \mathbf{d}_j \right), \\ \frac{\tilde{q}_i}{1 - \tilde{q}_i} &= \frac{q_i}{1 - q_i} \times \frac{\mathcal{CN}(0, \beta_i^{-1})}{\mathcal{CN}(\tilde{\mu}_i, \tilde{\beta}_i^{-1})}. \end{aligned}$$

To ensure a tractable posterior distribution $p(\beta_0 | \mathbf{a}, \mathbf{r}, b_n, d_n)$ of β_0 , it is also assumed that β_0 follows a Gamma distribution $\text{Gamma}(b_n, d_n)$. Therefore $p(\beta_0 | \mathbf{a}, \mathbf{r}, b_n, d_n)$ is easily given by

$$p(\beta_0 | \mathbf{a}, \mathbf{r}, b_n, d_n) = \text{Gamma}(b_n + L, d_n + \|\mathbf{r} - \mathbf{D}\mathbf{a}\|_2^2). \quad (19)$$

The estimate of the coefficient vector $\mathbf{a}$ is given empirically by averaging the last few samples of the Gibbs sampler. The successive steps of Gibbs sampler for our problem are summarized in Algorithm 1. The convergence of the MCMC sampling is generally hard to be diagnosed. However, empirical experiments show that the algorithm will converge after several dozens of iterations, which will be shown later.

Algorithm 1 Gibbs sampler

Initialize the maximum iteration number $N_{Maxiter}$, iteration counter $l = 0$ and initial value of parameters $\{\beta, \mathbf{q}, \beta_0, b, d, b_n, d_n, e_k, f_k\}$.

while $l \leq N_{Maxiter}$ **do**

for $i = 1 : MN$ **do**

 Sample a_i according to $p(a_i | \mathbf{a} \setminus a_i, \mathbf{r}, \mathbf{q}, \beta, \beta_0)$

 in (18);

 Sample β_i according to $p(\beta_{m,n} | \mathbf{a}, b, d)$ in (11),

 (12) or (13) corresponding to different continuity patterns;

 Sample q_i according to $p(q_{m,n}^k | \mathbf{a}, e_k, f_k)$

 in (14) corresponding to different sparsity patterns;

end for

Sample β_0 according to $p(\beta_0 | \mathbf{a}, \mathbf{r}, b_n, d_n)$ in (19).

end while

$\hat{\mathbf{a}}$ and $\hat{\beta}_0$ are to be obtained by the empirical mean of the samples of $\mathbf{a}$ and β_0 in the last few iterations, respectively.

B. Discussion on the Proposed Imaging Algorithm

Comparing to the previous work on ISAR and SAR imaging based on compressive sensing [1]-[8], the main difference and contribution of this paper is to exploit the continuity structure of the target scene in the sparse recovery method. Since the continuity structure is imposed statistically on the Bayesian prior instead of directly on the signal vector, the proposed method is flexible enough to perform successfully in the scenario of relatively weak continuity. Comparing to the

l_1 -norm based convex optimization approaches, the proposed method can achieve high-resolution in the scenarios of low SNR signals with only a few measurements, and is free of tedious parameter selection since the method itself is a data-driven learning process. However, there are still some limitations:

1) *Dependence on the basic assumption*: The performance of the proposed imaging algorithm depends on the correctness of the basic assumption that the scatterers of the target are located continuously in the target scene. The larger the deviation of the true target scene from the assumption, the worse the algorithm will perform especially under heavy noise or small sampling ratio situation. The limitation due to weak continuity is demonstrated and explained in Fig. 12 (h) in the subsection V-D. Remarkably to notice, however, this dependence limitation will not lead to degraded performance when the signal has a high SNR, since the continuity structure is imposed statistically on the Bayesian prior. This argument has been validated empirically in Fig. 12 (d) where the target scene is successfully recovered.

2) *Computation burden*: Since the inference of model parameters is carried out by the Gibbs sampler which inherently requires sequential sampling and several dozens of iterations, the computational burden is much heavier than that of the conventional RDA. It is easy to show from Algorithm 1 that the computational complexity of the proposed method is $O(L \times MN \times N_{Maxiter})$. For comparison, the computational complexities for the BP used in most of sparse representation based ISAR [2][8] and SAR [23][24] imaging methods, the BCS used in [6] and the proposed method are summarized in Table I. As demonstrated later, the proposed method converges quickly, $N_{Maxiter}$ used in our method is much smaller than L , i.e., $N_{Maxiter} \ll L$. Therefore, the computational complexity of the proposed algorithm can be written as $O(L \times MN)$, and is smaller than that of the BP. The computational complexity of the BCS is similar to that of the BP, if $N_{Maxiter} \approx \frac{MN}{L}$. However, to guarantee a good recovered result, L cannot be much smaller than MN . Therefore, we generally have $N_{Maxiter} \gg \frac{MN}{L}$ and the computational complexity of the BCS is relatively larger than that of the BP. Although the computational complexity of the method is relatively smaller than that of the BP and the BCS, it is still much larger than that of the RDA.

3) *Other limitations*: Since the formulated imaging problem is under-determined, the proposed imaging algorithm cannot deal with complex target scene where scatterers are extremely dense. Another limitation is that the algorithm proposed currently has no ability of de-speckling which may be critical for certain applications.

V. EXPERIMENTS AND PERFORMANCE COMPARISONS

In this section, we first analyze the behavior of the maximal coherence of dictionary $\mathbf{D}$ versus the number of the random samples in the range domain. Then, synthetic and practical data experiment results of the ISAR images produced by the proposed algorithm and others existing methods are studied and compared in this section to show the merits of our

proposed algorithm. At the same time, the performance of the proposed method with respect to the target sparsity is evaluated. The parameters of the proposed algorithm used in the experiments are given as follows: $b = d = b_n = d_n = 10^{-6}$, $(e_0, f_0) = (1/L, 1 - 1/L)$, $(e_1, f_1) = (1/L, 1/L)$, $(e_2, f_2) = (1 - 1/L, 1/L)$, the initial value of β is set to be $1/(\mathbf{D}^H \mathbf{r})$, $\mathbf{q} = \mathbf{1}$, $\beta_0 = 1/\text{var}(\mathbf{r}) \times 10^{-2}$, the maximum iteration number in the Gibbs sampler $N_{Maxiter}$ is 50 and the mean of the last 5 samples are computed as the estimate of $\mathbf{a}$.

A. Dictionary Coherence Evaluation

Before conducting the experiments, the behavior of the maximal mutual coherence of the dictionary $\mathbf{D}$ versus the number of random samples is investigated firstly by Monte Carlo method with 100 independent trials. The maximal mutual coherence of $\mathbf{D}$ versus the number of random samples is evaluated when $B/F_c = 0.15, 0.2$ and $N = 50$. The choices of $B/F_c = 0.15$ and $N = 50$ are to be consistent with the parameters used in section V-B. The choice of $B/F_c = 0.2$ is to show that dictionary with a low maximal mutual coherence can be achieved by designing the radar system with proper values of B and F_c . The maximal mutual coherence of the random Gaussian dictionary is also provided as a comparison. As shown in Fig. 4(a), when $B/F_c = 0.15$, maximal mutual coherence of the $\mathbf{D}$ is similar to that of the random Gaussian dictionary when 20 measurements in the range domain are used. However, when the number of measurements decreases, the maximal mutual coherence becomes smaller than that of the random Gaussian dictionary. As shown both in Fig. 4(a) and (b), the maximal mutual coherence of the dictionary increases as the number of the measurements decreases. The maximal coherence of dictionary $\mathbf{D}$ is always smaller than that of the random Gaussian dictionary when $B/F_c = 0.2$ as shown in Fig. 4(b). It is concluded that the dictionary is not worse than the random Gaussian matrix which has been found to satisfy the RIP with a high probability.

Fig. 4. Maximal mutual coherence versus the number of measurements (a) $B/F_c = 0.15$, (b) $B/F_c = 0.2$.

B. Synthetic Experiments

To favor the continuity patterns exploited in this paper, we assume that scatterers are on-grid and locate continuously in range and cross-range domains in the target scene. Suppose the target is 500 km away from the Radar in the direction of the LOS and has no translational motion. The target rotates with a constant velocity of $\omega = 0.1$ rad/s. In total, $N = 50$

TABLE I
COMPARISON OF THE COMPUTATIONAL COMPLEXITY

Algorithm	Proposed	BP [37]	BCS [39]
Complexity	$O(L \times MN \times N_{Maxiter})$	$O(L^2 \times MN)$	$O(L^3 \times N_{Maxiter})$

TABLE II
SYSTEM PARAMETERS FOR THE SIMULATED DATA

Carrier frequency F_c	10 GHz
Bandwidth B	1.5 GHz
Range resolution $R_r = C/(2B)$	0.1 m
Pulse repetition frequency (PRF)	100 Hz
Cross-range resolution D_r	0.3 m

pulses are collected. The radar system parameters are given in Table. II.

Since 50 pulses within a dwell time of 0.5s are used, the total rotated angle θ is 2.9° . Therefore, the cross-range resolution $\frac{\lambda}{2\theta}$ is 0.3 m. Suppose the target scene of interest is of size $4 \times 15 m^2$ with a resolution of $0.1 \times 0.3 m^2$. The scatterers are present in the center of the target scene and their locations are shown in Fig. 5, where 57 scatterers are present and located continuously along both range and cross-range domains. The amplitudes of the scatterers follow a complex Gaussian distribution with a mean of 2.0 and a variance of 0.5. The received signal is corrupted by the complex white Gaussian noise with a power of 0.2. Measurements are obtained by sampling randomly in the range domain and measurements of one half, one fourth and one eighth of the full data set, i.e., 20, 10 and 5 measurements are used in the following synthetic experiments. The results of the RDA [16], BP as used in [2] [8][23] and the Bayesian compressive sensing (BCS)³ [6][18] are also provided for comparison.

Fig. 5. The scatterer locations

Fig. 6 gives the obtained ISAR images from different methods with one half of the full measurements. The noiseless target scene is shown in Fig. 6(a) where the scatterers are clustered together showing a strong continuity. The image by RDA with full measurements is given in Fig. 6(b) where the target image is clear with slight background noise. However, when one half of the measurements in the range frequency domain (randomly selected) are used by RDA, the image becomes blurred by the increased background noise shown in Fig. 6 (c). The reconstructed images with SPF by BP, BCS and the proposed method are given in Fig. 6(d)-(f). The result of BP in Fig. 6(d) is much better than that of the RDA, but is subjected to a careful selection of the regularization parameter

Fig. 6. ISAR images achieved using SPF by different methods: (a) the target scene, (b) the RDA with full measurements; (c) the RDA; (d) the BP; (e) the BCS; (f) the proposed method.

for the algorithm. Both BCS and the proposed method are free of parameter selection. However, BP does not perform well, as shown in Fig. 6(e), because of its difficulties in obtaining satisfactory image by using a small number of the measurements. The diminishing of some scatterers in Fig. 6 (e) may be caused by the noise suppression process of the BCS to keep a relatively sparse and clear target scene. It is observed that some of the scatterers with low amplitudes are eliminated by the BP in Fig. 6(d). In contrast, since the proposed method exploits the structure of the signal and enforces the continuity of the scatterers, the scatterers with relatively low amplitudes can be preserved. As shown in Fig. 6(f), every scatterer in the target scene is preserved and the background noise is effectively removed.

Fig. 7 shows the performances of different methods when the number of the measurements decreases to 1/4 and 1/8 of the original one. Since BCS cannot work under those settings, we only provide the results obtained by RDA, BP and the proposed method. As shown in Fig. 7(a) and (b), the RDA method shows worsen performance as the number of the measurements used decreases and fails when 1/8 of measurements is used. Both BP and the proposed method with SPFs using 1/4 of the full measurements perform well as shown in Fig. 7(c) and (e), where the proposed method outperforms BP slightly by removing the weak artificial points

³Available at website: <http://people.ee.duke.edu/~lcarin/BCS.html>

Fig. 7. ISAR image comparison in terms of the number of SPF. (a) The RDA using 1/4 of measurements; (b) The RDA using 1/8 of measurements; (c) The BP using 1/4 of measurements; (d) The BP using 1/8 measurements; (e) The proposed method using 1/4 measurements; (f) The proposed method using 1/8 measurements.

around the target while preserving some weak scatterers. When the number of the measurements decreases to 1/8 of the original one, the result of BP is unsatisfactory in Fig. 7(d) where more than 2/3 of the scatterers are missing with many strong artificial points around. In comparison, the result of the proposed method is much better as shown in Fig. 7(f) where most of the scatterers are preserved with fewer artificial points.

The convergence curve of the Gibbs sampler based on the simulated data is given in Fig. 8 with 10 independent experiments. The normalized reconstruction error is used to diagnose the convergence of the method. For the ISAR image, since we are more interested in the absolute amplitude of the scatterer than the complex one, the normalized reconstruction error of the absolute amplitude vector $\bar{\mathbf{a}} = [|a_{1,1}| \cdots |a_{M,N}|]^T$ is computed by:

$$\left\| \frac{\bar{\mathbf{a}}}{\|\bar{\mathbf{a}}\|_2} - \frac{\hat{\mathbf{a}}}{\|\hat{\mathbf{a}}\|_2} \right\|_2 \quad (20)$$

where $\hat{\mathbf{a}}$ is the absolute sample value of the sparse coefficient in each iteration of the Gibbs sampler. The noise power added in each experiment is 0.2 and one half of the measurements, i.e., 20 measurements in the range frequency domain, are used. As shown in Fig. 8, the MCMC converges fast and the relative reconstruction error remains roughly unchanged after about 25 iterations. Therefore, a small number of iterations is sufficient to accurately represent the full posterior on all model parameters. Note that there is a remaining convergence error even after 30 iterations because of the mechanism of the Gibbs sampler and the random noise. Since the final estimate is the average of the last 5 samples of the Gibbs sampler, the error

of the final estimate is expected to be similar or smaller than the convergence error, which is also demonstrated in Fig. 9(a) that when 20 random measurements are used under a noise of power 0.2, the error of the final estimate given by the proposed method is 0.1772.

Fig. 8. The converge curve.

Since the RDA does not have the ability to suppress either the additive noise or the noise from undersampling, we simply compare the performance of the BP with that of the proposed algorithm in terms of normalized mean square error (NMSE). The NMSE is computed as the mean of the normalized reconstruction error defined in (20) from 50 independent experiments with $\hat{\mathbf{a}}$ replaced by the absolute value of the final estimate. Fig. 9(a) and (b) give the NMSE of the reconstructed images by the proposed algorithm and the BP method with respect to the number of measurements and the noise power, respectively. It is seen from Fig. 9(a) and (b), the proposed method achieves smaller NMSE than the BP method due to the exploitation of the continuity structure of the target scene. The CPU time consumed by the proposed algorithm

Fig. 9. The NMSEs: (a) against the number of measurements (noise power = 0.2), (b) against the noise power with 10 measurements in the range domain.

and the BP method against the number of measurements are listed in Table III for comparison, in which each value is evaluated based on the average of 5 trials. In each trial, measurements of the same number are randomly selected from the range domain. We use CVX toolbox⁴ to implement the BP method. Note that 50 iterations are performed for the Gibbs sampler. Parameters are the same as those used for Fig. 9(a). It is observed from Table III that the proposed method is much faster than the BP method, especially when the number of measurements used is relatively large. It has been shown in section IV-B that the computational complexity of the proposed method is $O(L \times MN \times N_{Maxiter})$ which increases linearly with the number of the measurements, as verified in Table III. However, the computational time for

⁴Available at website: <http://cvxr.com/cvx/>

TABLE III
THE REQUIRED CPU TIME (S)

Measurements	5	10	15	20	25
Proposed	85.1	206.3	310.5	418.7	538.7
BP	164.7	616.9	1312.6	2423.9	3608.4

Fig. 11. The NMSE versus the degree of the target sparsity.

the BP method increases quadratically with the number of measurements when it is implemented by CVX toolbox. Note that the computational complexity of the BP method depends on its implementation. It is possible to speed up the BP method by other implementations such as alternating direction method of multipliers (ADMM), which is out of the scope of this paper.

C. Performance evaluation with respect to the target sparsity

To evaluate the performance of the proposed method with respect to the target sparsity, we design and conduct the following simulations. The parameters of radar system are the same as those shown in Table II. A ratio of 1/4 of the full measurements in the range domain is used. Other parameters remain the same as those used in Fig. 6, except the target scene. We increase successively the number of the nonzero scatterers, i.e., the degree of the sparsity, in the following simulations, and evaluate the NMSE of the results obtained by the proposed method. The results are given in the Fig. 10. The number of the nonzero scatterers is 7×7 , 11×11 , 13×13 and 15×15 for each column in Fig. 10. The degree of the target sparsity is computed and shown below each column along with the NMSE for each image obtained by the proposed method. The average NMSE obtained by 10 independent trails is shown in Fig. 11. As shown in Fig. 11, the performance of the proposed method decreases as the degree of the target sparsity increases. We emphasize that the degree of sparsity at which a correct target image can be recovered depends highly on the coherence of the dictionary. The lower the maximal coherence of the dictionary, the higher the sparsity degree can be allowed. Therefore, it is expected to achieve a higher degree of the target sparsity if we design a dictionary with a lower coherence.

D. Experiment with data set “Mig-25”

The proposed algorithm is also tested by a widely used synthetic data set “Mig-25” downloaded from the homepage of V. C. Chen⁵. The Radar system model parameters are given in Table IV. The number of points in the range domain is 64 and

TABLE IV
SYSTEM MODEL PARAMETERS FOR DATA “MIG-25”

Carrier frequency F_c	9 GHz
Bandwidth B	512 MHz
Range resolution $R_r = C/(2B)$	0.293 m
Pulse repetition frequency (PRF)	15 KHz

32 pulses are collected. The additive noise follows a complex white Gaussian distribution with a variance of 5. The signal to noise ratio (SNR) in this case equals to 7.5dB. The results obtained by different methods using different proportions of measurements are given in Fig. 12. Note that the results in Figs. 6 and 7 are presented in the axes of range and cross-range with the unit of meters since the scatterer positions can be easily calculated by the known rotating velocity of the target. However, since the information of the rotating velocity is unavailable for “Mig-25” and the data “Yak-42” used in the following subsections, the results in the rest of the paper will be presented with axes labelled by “Range cell” and “Doppler cell”.

The result of the RDA with full measurements in Fig. 12(a) shows some continuity in the range domain and a weak continuity in the Doppler (cross-range) domain of the simulated target. The RDA method performs well in noisy environment, as shown in Fig. 12(e). However, the performance severely degrades when 1/2 and 1/4 measurements are used, as shown in Fig. 12(b) and (f), especially in Fig. 12(f) where target cannot be observed. The images given by the BP and the proposed method are much better, as shown in Fig. 12(c) and (d) using 1/2 measurements. In comparison of Fig. 12(c) and (d) with Fig. 12(a), it seems that the result of the proposed method is slightly better than that of BP in term of removing the artificial points outside the target region and preserving the weak scatterers. Neither of the two methods can produce an acceptable result with 1/4 of the measurements. But a plane profile can be seen from the image obtained by the proposed method in Fig. 12. However, the noise components in the target region are preserved and are taken wrongly as the signal components by the proposed algorithm since it always enforces a continuity inside the target region. Comparing to Fig. 12(h), the noise components inside the target region in Fig. 12(g) are efficiently removed by the BP which does not use any continuity assumption. However, we can hardly observe a plane from the image in Fig. 12(g) where only some strong target scatterers are preserved with many comparable artificial points around.

E. Practical Data Experiments with data set “Yak-42”

The data used in the experiments is “Yak-42” plane data set. The Radar system parameters are given in Table V.

TABLE V
SYSTEM PARAMETERS FOR DATA “YAK-42”

Carrier frequency F_c	10 GHz
Bandwidth B	400 MHz
Range resolution $R_r = C/(2B)$	0.375 m
Pulse repetition frequency (PRF)	25 Hz

⁵Available at website: <http://airborne.nrl.navy.mil/~vchen/tftsa.html>

Fig. 10. Performance of the proposed method versus the degree of the target sparsity. The first row presents the target scenes with increased degree of target sparsity; the second row shows the images obtained by the RDA; the third row presents the images obtained by the proposed method; and the last row indicates the degree of the target sparsity and the NMSE of the results obtained by the proposed method.

Fig. 12. Comparison of ISAR images against number of SPFs achieved by the (a) RDA with full measurements; (b) RDA using 1/2 of measurements; (c) BP using 1/2 of measurements; (d) proposed method using 1/2 of measurements; (e) RDA with full measurements ($SNR = 7.5\text{dB}$); (f) RDA using 1/4 of measurements; (g) BP using 1/4 of measurements; (h) proposed method using 1/4 of measurements.

The number of points in the range domain is 128 and the number of pulses is 32 within a dwell time of 1.28s. The RD image with full measurements are given in Fig. 13 where a relatively strong continuity of the plane fuselage and a weak continuity around the head and wings of the plane can be observed. Since the data size of “Yak-42” is relatively large, we only perform 35 Gibbs sampling iterations and the estimate of $\mathbf{a}$ is given by the empirical mean of the samples in the last 5 iterations.

The results of “Yak-42” plane by different methods are given in Figs. 14 and 15. In Fig. 14, the results obtained by different methods with various numbers of SPFs are compared. The effect of under-sampling becomes extremely severe in

Fig. 14(a) and (b). In Fig. 14(b), the head and wings of the target are nearly covered by the noise. The target images given by both the BP and proposed method using one half of the measurements are satisfactory. In the image obtained by the BP in Fig. 14(c), some weak scatterers on the target region are missing and several weak artificial points are around. The image obtained by the proposed method in Fig. 14(e) is relatively better, where the artificial points outside the target region are almost removed and the continuity property of the target image is well preserved. The performances of both the proposed method and the BP using 1/4 of measurements degrade. Some scatterers on the target are removed along with the noise by both the BP and the proposed method. The part

Fig. 13. The ISAR image of Yak-42 airplane obtained by the RDA with full measurements.

Fig. 14. ISAR images versus the number of SPFs obtained by the (a) RDA using 1/2 of measurements; (b) RDA using 1/4 of measurements; (c) BP using 1/2 of measurements; (d) BP using 1/4 of measurements; (e) proposed method using 1/2 of measurements; (f) proposed method using 1/4 of measurements.

connecting the fuselage and the head of the target is lost in Fig. 14(f). The image given by BP in Fig. 14(d) seems to be too sparse where too many scatterers on the target region are missing comparing to that shown in Fig. 14(f).

Figure 15 compares the results of different methods using one half of the measurements with noise powers of 150 and 300, respectively. The corresponding SNRs are 11.3dB and 5.3dB, respectively. The results of the RDA with full measurements are shown in Fig. 15(a) and (e), respectively. The plane profile can still be clearly recognized. However, when one half of the measurements are used, the head and wings of the plane can hardly be seen from the results of RDA in Fig. 15(b) and (f). In Fig. 15(f), only a part of the fuselage is seen. Although the result of the BP with a noise power of 150 shown in Fig. 15(c) preserves most of the strong scatterers, the artificial points are still present on the entire target scene which are comparable to those of the weak scatterers. The

result of the proposed method in Fig. 15(d) seems to be better with a clean background, although some scatterers on the head and the lower wing are missing. The result of BP with noise power of 300 becomes poor, as shown in Fig. 15(g) where only several strong scatterers can be seen and others are immersed in the remaining artificial points. However, the proposed method can still work and provide in Fig. 15(h) a rough but clear profile of the plane with the continuity among parts of the plane preserved.

In addition to those visual result, we further use image entropy and the correlation value [40] with the reference image as two criteria to quantitatively evaluate the quality of the recovered images after random sampling. In the correlation criterion, the gray level image obtained by the RDA using the original dataset with full measurements is used as the reference image denoted by $\mathbf{X}_o$. Suppose the obtained image by different methods is denoted as $\mathbf{X}_r$, the correlation is then defined as

$$\text{Corr}(\mathbf{X}_r, \mathbf{X}_o) = \frac{|\langle \text{vec}(\mathbf{X}_r), \text{vec}(\mathbf{X}_o) \rangle|}{\|\text{vec}(\mathbf{X}_r)\|_2 \|\text{vec}(\mathbf{X}_o)\|_2},$$

which evaluates the similarity between the recovered image and the reference image, where $\text{vec}(\mathbf{X})$ denotes the function which vectorizes the matrix $\mathbf{X}$ and $\langle \cdot, \cdot \rangle$ denotes the inner product. The higher the correlation value a method can provides, the better the method preserves the information of the target. The entropy of the image, given by $\text{Entropy}(\mathbf{X}_r) = -\sum \mathbf{p} \cdot \log 2(\mathbf{p})$, where $\mathbf{p}$ is the histogram of the recovered gray level image, measures the focusing quality of the obtained ISAR image. The more the image is focused, the smaller its entropy is. The correlation and the entropy values of the obtained images by different methods under different SNRs and sampling ratio are summarized in Table VI. As indicated in Table VI, the results obtained by the proposed method maintain a higher correlation with the reference image and lower image entropy than those given by the BP.

F. Discussion on the Results

The proposed method shows a great potential in removing the noise outside the target region based on the assumption that target scatterers are continuously located in a particular region while the randomly located noise can hardly form a cluster. As long as the target scene fits the assumption well, the proposed method will greatly outperform those imaging algorithms based on conventional sparsity assumption, which is validated by the results shown in Figs. 6 and 7. If weak continuity of the target occurs, such as the continuity exhibited in dataset “Mig-25”, and relatively heavy noise components inside target region due to undersampling, the proposed method tends to give a profile of the target as shown in Fig. 12(h), where detailed information inside the target region is lost. Comparing to the ISAR image achieved by BP, another advantage of the proposed method is to be free of regularization parameter selection. All the parameters and hidden variables including the noise power can be learned directly from the data.

VI. CONCLUSION

The conventional CS-based ISAR imaging methods only assume that scatterers on target scene are sparse or compress-

Fig. 15. ISAR images comparison with different noise powers. (a) Image achieved by RDA under noise (noise power=150); (b) Image achieved by RDA (noise power=300); (c) Image achieved by RDA by using 1/2 of measurements (noise power=150); (d) Image achieved by RDA using 1/2 of measurements (noise power=300); (e) Image achieved by BP by using 1/2 of measurements (noise power=150); (f) Image achieved by BP using 1/2 of measurements (noise power=300); (g) Image achieved by proposed method using 1/2 of measurements (noise power=150); (h) Image achieved by proposed method using 1/2 of measurements (noise power=300).

TABLE VI
PERFORMANCE EVALUATION BY **Corr** AND **Entropy**

SNR (dB)	10				5	
Sampling ratio	1/2		1/4		1/2	
Algorithm	BP	Proposed	BP	Proposed	BP	Proposed
<i>Corr</i>	0.8139	0.8567	0.7579	0.8159	0.7689	0.8239
<i>Entropy</i>	1.9326	0.2456	1.1568	0.1848	1.7106	0.2307

ible without considering any particular signal structure. In contrast, a new ISAR imaging method within the Bayesian framework is proposed in this paper to obtain an enhanced target image by exploiting the structural continuity of the scatterers on the target scene. To impose the sparsity on the signal, the amplitude of each scatterer is assumed to follow a Bernoulli-Gaussian distribution. Then a fully correlated prior is imposed to further enforce the continuity by assuming that scatterers follow certain continuity pattern. By exploiting the structural continuity of target scene, the proposed algorithm requires fewer measurements in range domain, which results in a simplified radar system with only a small number of random sparse probing frequencies, and achieves effective denoising outside target region and higher resolution at the expense of relatively heavy computational complexity.

Due to the computational complexity, the proposed algorithm seems to be less applicable to the SAR imagery which involves a huge amount of data processing since usually an extensive target scene is of interest in SAR. An efficient variational inference which exploits the structure of the signal for large-scale Bayesian compressive sensing will be investigated in future research for alleviating the computational complexity. Since speckle noise widely exists in SAR/ISAR, we are to investigate the possibility of Bayesian compressive sensing technique to deal with it.

APPENDIX

The approximate maximal coherence of the dictionary D with full measurements is derived in this Appendix. When full measurements are used, the number of measurements $L = MN$. The mutual coherence between atom $\mathbf{d}_k$ and $\mathbf{d}_{k'}$, denoted as $h(k, k')$, where $k \neq k'$, can be computed as

$$h(k, k') = \frac{|\mathbf{d}_k^H \mathbf{d}_{k'}|}{MN} \quad (21)$$

Substituting (6) into (21), we have

$$h(k, k') = \frac{1}{MN} \left| \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \exp \left\{ -j \frac{4\pi f_m}{c} \left((y_{2,k} - y_{2,k'}) + \frac{(y_{1,k} - y_{1,k'}) |\Omega_{eff}|}{PRF} n \right) \right\} \right|. \quad (22)$$

Let $\Delta y_2 = (y_{2,k} - y_{2,k'})$ and $\Delta y_1 = \frac{(y_{1,k} - y_{1,k'}) |\Omega_{eff}|}{PRF}$, it can be easily found that

$$\begin{aligned} \Delta y_2 &= l_2 \frac{c}{2B}, & -M+1 &\leq l_2 \leq M-1, \\ \Delta y_1 &= l_1 \frac{\lambda}{2N}, & -N+1 &\leq l_1 \leq N-1. \end{aligned}$$

TABLE VII
THE MAXIMAL COHERENCE

B/F_c	0.04	0.06	0.08	0.1	0.12	0.14	0.15	0.16	0.18	0.2
$N = 50$	0.2738	0.2991	0.2848	0.2118	0.1875	0.1576	0.1422	0.1382	0.1287	0.1096
$N = 100$	0.2800	0.1868	0.1383	0.1088	0.0909	0.0780	0.0731	0.0700	0.0647	0.0598

Substituting the above equations into (22), after some direct derivations, we have

$$\begin{aligned}
& h(k, k') \\
&= \frac{1}{MN} \left| \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \exp \left\{ -j \frac{4\pi f_m}{c} \left(l_2 \frac{c}{2B} + l_1 \frac{\lambda}{2N} n \right) \right\} \right| \\
&= \frac{1}{MN} \left| \sum_{n=0}^{N-1} \exp \left\{ -j \frac{4\pi}{c} (F_c - 2\Delta f) \frac{\lambda}{2N} l_1 n \right\} \right. \\
&\quad \times \left. \frac{\sin \left(\frac{2\pi B}{c} \left(l_2 \frac{c}{2B} + l_1 \frac{\lambda}{2N} n \right) \right)}{\sin \left(\frac{2\pi \Delta f}{c} \left(l_2 \frac{c}{2B} + l_1 \frac{\lambda}{2N} n \right) \right)} \right| \quad (23)
\end{aligned}$$

where $\Delta f = B/M$. If M is large enough, that is Δf is extremely small, then the last item in (23) can be simply approximated by a sinc function, and $2\Delta f$ in the first term in (23) can be ignored. Therefore, (23) can be written as

$$\begin{aligned}
h(k, k') &\approx \frac{1}{N} \left| \sum_{n=0}^{N-1} \exp \left\{ -j 2\pi \frac{l_1}{N} n \right\} \right. \\
&\quad \times \left. \text{sinc} \left(\pi \left(l_2 + \frac{B}{F_c} \frac{l_1}{N} n \right) \right) \right|. \quad (24)
\end{aligned}$$

As seen from (24), the coherence is determined by N , B/F_c , l_1 and l_2 . In the following, the effect of l_1 and l_2 on the value of h is discussed.

- 1) When $l_1 = l_2 = 0$, h evaluates the coherence of one atom with itself, and $h = 1$.
- 2) When $l_1 = 0$ and $l_2 \neq 0$, we have

$$h \approx \frac{1}{N} \left| \sum_{n=0}^{N-1} \text{sinc}(\pi l_2) \right| = 0$$

- 3) When $l_1 \neq 0$, the analyses of h become complicated. Given a fixed N , h highly depends on the value of B/F_c .

Generally, signal of X-band is used for ISAR imaging with F_c around 8–10GHz and B around 500–1000MHz. Therefore, we consider in this paper $0.04 \leq \frac{B}{F_c} \leq 0.2$. Since close form of h over B/F_c is difficult to be derived, we present maximal value of h for two different atoms numerically in Table VII when $N = 50$ and 100. As indicated by Table VII, the maximal coherence decreases with the ratio B/F_c in the range 0.08–0.2. And the maximal coherence for $N = 100$ is almost one half of that for $N = 50$ when $0.08 \leq \frac{B}{F_c} \leq 0.2$. Therefore, we can always design $B/F_c = 0.2$ for radar system to guarantee a low maximal coherence in dictionary $\mathbf{D}$.

REFERENCES

- [1] V. M. Patel, G. R. Easley, D. M. Healy, Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 244-254, Apr. 2010.
- [2] L. Zhang, M. Xing, C. Qiu, J. Li, J. Sheng, Y. Li, and Z. Bao, "Resolution enhancement for inversed synthetic aperture radar imaging under low SNR via improved compressive sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3824-3838, Oct. 2010.
- [3] G. Li, H. Zhang, X. Wang and X. Xia, "ISAR 2-D imaging of uniformly rotating targets via matching pursuit," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1838-1846, Apr. 2012.
- [4] G. Xu, M. Xing, L. Zhang, Y. Liu, and Yachao Li, "Bayesian inverse synthetic aperture radar imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 6, pp. 1150-1154, Nov. 2011.
- [5] L. Zhao, G. Bi, L. Wang and H. Zhang, "An Improved Auto-Calibration Algorithm Based on Sparse Bayesian Learning Framework," *IEEE Signal Process. Lett.*, vol. 20, no. 9, pp. 889-892, Sep. 2013.
- [6] J. Xu Y. Pi and Z. Cao, "Bayesian compressive sensing in synthetic aperture radar imaging," *IET Radar, Sonar and Navigation*, vol. 6, no. 1, pp. 2-8, Jan. 2012.
- [7] G. Zhao, Z. Wang, Q. Wang, G. Shi and F. Shen, "Robust ISAR imaging based on compressive sensing from noisy measurements," *Signal Process.*, vol. 92, no. 1, pp. 120-129, Jan. 2012.
- [8] H. Wang, Y. Quan, M. Xing, and S. Zhang, "ISAR imaging via sparse probing frequencies," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 451-455, May 2011.
- [9] I. F. Gorodnitsky, and B. D. Rao, "Sparse signal reconstruction from limited data using FOCUSS : a re-weighted minimum norm algorithm," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 600-616, Mar. 1997.
- [10] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129-159, Mar. 2001.
- [11] R. Tibshirani, "Regression shrinkage and selection via the lasso", *J. R. Statist. Soc. vol. B*, no. 58, pp. 267-288, 1996.
- [12] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397-3415, Dec. 1993.
- [13] J. A. Tropp, A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.
- [14] D. Needell, R. Vershynin, "Uniform uncertainty principle and signal recovery via regularized orthogonal matching pursuit," *Found. Comput. Math.*, vol. 9, pp. 317-334, 2009.
- [15] D. L. Donoho, Y. Tsaig, I. Drori and J-L Starck, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1094-1121, Feb. 2012.
- [16] J. L. Walker, "Range-Doppler imaging of rotating objects," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 16, no. 1, pp. 23-52, Jan. 1980.
- [17] F. Berizzi, E. Dalle Mese, M. Diani and M. Martorella, "High-resolution ISAR imaging of maneuvering targets by means of the range instantaneous Doppler technique: modeling and performance analysis," *IEEE Trans. Image Process.*, vol. 10, no. 12, pp. 1880-1890, Dec. 2001.
- [18] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346-2356, June 2008.
- [19] M. J. Beal, "Variational algorithms for approximate Bayesian inference," Ph.D. dissertation, Gatsby Computational Neuroscience Unit, Univ. College London, London, U.K., 2003.
- [20] L. He, H. Chen, and L. Carin, "Tree-Structured Compressive Sensing With Variational Bayesian Analysis," *IEEE Signal Process. Lett.*, vol. 17, no. 3, pp. 233-236, Mar. 2010.
- [21] L. He and L. Carin, "Exploiting structure in wavelet-based Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3488-3497, Sep. 2009.
- [22] M. Çetin, and W. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Trans. Image Process.*, vol. 10, no. 4, pp. 623-631, 2001.
- [23] M. Tello Alonso, P. Lopez-Dekker, and J.J. Mallorqui, "A Novel Strategy for Radar Imaging Based on Compressive Sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4285-4295, 2010.
- [24] N. Onhon, and M. Cetin, "A Sparsity-Driven Approach for Joint SAR Imaging and Phase Error Correction," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2075-2088, 2012.

- [25] M. Zhou, H. Chen, J. Paisley, L. Ren, L. Li, Z. Xing, D. Dunson, G. Sapiro and L. Carin, "Nonparametric Bayesian dictionary learning for analysis of noisy and incomplete images," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 130 - 144, Jan. 2012.
- [26] L. Yu, H. Sun, J. P. Barbot, and G. Zheng, "Bayesian compressive sensing for cluster structured sparse signals," *Signal Process.*, vol. 92, no. 1, pp. 259-269, Jan. 2012.
- [27] N. Dobigeon, A. Basarab, D. Kouame and J. Tournet, "Regularized Bayesian compressed sensing in ultrasound imaging," in *Proc. IEEE 20th European Signal Process. Conf. (EUSIPCO)*, Bucharest, Romania, Aug. 2012, pp. 2600-2604.
- [28] R.G. Baraniuk, V. Cevher, M.F. Duarte and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982-2001, Apr. 2010.
- [29] Y. C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042-3054, June 2010.
- [30] X. Lv, G. Bi, and C. Wan, "The group lasso for stable recovery of block-sparse signal representations," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1371-1382, Apr. 2011.
- [31] L. Wang, G. Bi, C. Wan and X. Lv, "Improved stability conditions of BOGA for noisy block-sparse signals," *Signal Process.*, vol. 91, no. 11, pp. 2567-2574, Nov. 2011.
- [32] D. Baron, M. F. Duarte, S. Sarvotham, M. B. Wakin, and R. G. Baraniuk, "Distributed compressed sensing," 2005, [online] Available: <http://dsp.rice.edu/cs/DCS112005.pdf>.
- [33] C. La and M. N. Do, "Tree-based orthogonal matching pursuit algorithm for signal reconstruction," in *IEEE Int. Conf. on Image Process. (ICIP)*, (Atlanta, GA), pp. 1277-1280, Oct. 2006.
- [34] M. F. Duarte, M. B. Wakin, and R. G. Baraniuk, "Wavelet-domain compressive signal reconstruction using a hidden Markov tree model," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, Las Vegas, NV, April 2008, pp. 5137-5140.
- [35] C. Quinsac, N. Dobigeon, A. Basarab, J. Y. Tournet, and D. Kouame, "Bayesian compressed sensing in ultrasound imaging," in *Proc. IEEE Int. Workshop Comput. Adv. in Multi-Sensor Adaptive Process. (CAMSAP)*, San Juan, Puerto Rico, Dec. 2011, pp. 101-104.
- [36] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, 2nd ed. New York: Springer, 2004.
- [37] P. Bühlmann, and L. Meier, "Discussion: One-step sparse estimates in nonconcave penalized likelihood models," *The Annals of Statistics*, vol. 36, no. 4, pp. 1534-1541, 2008.
- [38] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489-509, Feb. 2006.
- [39] D.P. Wipf, and B.D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153-2164, Aug. 2004.
- [40] L. Yang, M. Xing, L. Zhang, J. Sheng, and Z. Bao, "Entropy-based motion error correction for high-resolution spotlight sar imagery," *IET Radar, Sonar Navigation*, vol. 6, no. 7, pp. 627-637, 2012.

Lu Wang received the B.Eng. in electrical and electronic engineering from Xidian University in 2007 and M.Eng. degree in signal processing from National Key Lab of Radar Signal Processing (RSP), Xidian University, China, in 2010. And she is currently pursuing the Ph.D. degree in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Her major research interests include sparse signal processing, radar imaging (SAR/ISAR), sonar signal processing.

Lifan Zhao received the B.Eng. degree in electronic engineering from Xidian University, Xi'an, China, in 2010. He is currently working towards the Ph.D. degree in electrical and electronic engineering from Nanyang Technological University, Singapore. His research interests include sparse signal recovery techniques and their applications.

Guoan Bi received the B.Sc. degree in radio communications from Dalian University of Technology, P. R. China, 1982, and the M.Sc. degree in telecommunication systems, and the Ph.D. degree in electronics systems, both from Essex University, U.K., in 1985 and 1988, respectively.

Since 1991, he has been with the school of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His current research interests include DSP algorithms and hardware structures, and signal processing for various applications including sonar, radar, and communications.

Chunru Wan received the B.Eng. and M.Eng. degrees in electrical and electronic engineering from Northwestern Polytechnical University, Xian, China, respectively, in 1985 and 1988. He received the Ph.D. degree in computer studies from Loughborough University, U.K., in 1996.

From 1988 to 1990, he was a researcher with the Institute of Acoustic Engineering, College of Marine Engineering, Northwestern Polytechnical University. From 1990 to 1992, he was with the Department of Electrical and Electronic Engineering, Loughborough University. From 1993 to 1998, he was with DSO National Laboratories as a Senior Engineer. Since 1998, he has been a faculty member with the School of Electrical and Electronic Engineering, Nanyang Technological University, while remaining an Adjunct Principal Member of Technical Staff with DSO National Laboratories. He has published more than 100 papers in international journals and conferences. His research interests are in signal processing, sonars, communications, parallel computing, underwater acoustics, and computational mathematics.

Lei Yang received the B.Eng. degree in electrical and electronic engineering from Xidian University, China, in 2007. He received the Ph.D. degree in signal processing from National Key Lab of Radar Signal Processing (RSP), Xidian University, in 2012. His major research interests include radar imaging (SAR/ISAR), interferometric SAR, and ground moving target indication. From 2012, he is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore.