Pattern-Coupled Sparse Bayesian Learning for Inverse Synthetic Aperture Radar Imaging

Huiping Duan, Lizao Zhang, Jun Fang, Lei Huang, and Hongbin Li, *Senior Member, IEEE*

***Abstract*—We propose a pattern-coupled sparse Bayesian learning method for inverse synthetic aperture radar (ISAR) imaging by exploiting a block-sparse structure inherent in ISAR target images. A two-dimensional pattern-coupled hierarchical Gaussian prior is proposed to model the pattern dependencies among neighboring scatterers on the target scene. An expectation-maximization (EM) algorithm is developed to infer the maximum a posterior (MAP) estimate of the hyperparameters, along with the posterior distribution of the sparse signal. Numerical results are provided to illustrate the effectiveness of the proposed algorithm.**

***Index Terms*—Block-sparse structure, expectation-maximization (EM), ISAR, pattern-coupled sparse bayesian learning.**

I. Introduction

THE application of sparse representation to SAR/ISAR imaging has attracted much attention over the past few years, e.g. [1]–[8]. This new class of advanced imaging methods present a number of unique advantages over conventional range-Doppler methods [9], including improved resolvability of point scatterers [1], [3], reduced speckle [4], and recoverability from fewer data samples [2]. The basic idea behind these works is to formulate SAR/ISAR imaging as a sparse signal recovery problem. In addition to the sparse structure, real-world SAR/ISAR images often have additional structures that can be utilized. For example, by exploiting the piecewise smoothness [4] and the continuity structure [6] of the target image, an enhanced robustness against speckle noise and a substantial performance improvement can be achieved. Also, in practice, the target of interest usually exhibits block-sparse structures in which nonzero large scatterers occur in clusters. This block-sparse pattern can be considered as a special form of the continuity structure where nonzero scatterers demonstrate continuity in both the range and cross-range domains.

This paper aims at leveraging the block-sparse structure of the target image to enhance ISAR imaging performance. We propose a two-dimensional pattern-coupled sparse Bayesian learning method which is a generalization of the conventional sparse Bayesian learning (SBL) method [10] to deal with block-sparse signals. The proposed method is effective and flexible to exploit the underlying block-sparse structures, without requiring the prior knowledge of the block partition. Experimental results demonstrate that the proposed method is able to achieve a substantial performance improvement over existing algorithms, including the conventional SBL method [10], [11].

II. ISAR Imaging Model

Suppose the radar transmits a linear frequency-modulated signal

$$s(\tau) = \text{rect}(\tau/T_p)\exp\left\{j2\pi\left(f_c\tau + \frac{\gamma}{2}\tau^2\right)\right\} \tag{1}$$

where τ is the fast time, T_p is the pulse duration, and $\text{rect}(\cdot)$ represents the unit rectangular function, f_c is the carrier frequency, and γ is the chirp rate. Consider the signal reflected from a single scatterer located at position (u, v) in the cross-range and range domains. Assume that the phase errors have been compensated by autofocus algorithms, e.g. [12]. After range compression, the echoed signal is given as

$$\begin{aligned} &r(\tau, t) \\ &= x \cdot \text{rect}(t/T_a)\text{sinc}\left(T_p\gamma[\tau - 2R(t)/c]\right)\exp\left(-j4\pi R(t)/\lambda\right) \end{aligned} \tag{2}$$

where x denotes the scattering amplitude, t is the slow time, T_a is the coherent processing interval, $\text{sinc}(\cdot)$ is the sinc function, c denotes the speed of light, λ is the wavelength, and $R(t)$ is the instantaneous distance between the scatterer and the radar which can be approximated by $R(t) \approx R_0 + v + u\omega t$, where R_0 is the target coordinate origin, and ω denotes the rotational angular velocity (see Fig. 1). Substituting $R(t)$ into (2), we arrive at

$$\begin{aligned} r(\tau, t) \approx\ & x \cdot \text{rect}(t/T_a)\text{sinc}\left(T_p\gamma[\tau - 2(R_0 + v)/c]\right) \\ & \exp\left(-j4\pi(R_0 + v)/\lambda\right)\exp(-j2\pi f t) \end{aligned} \tag{3}$$

where $f \triangleq 2u\omega/\lambda$ is the Doppler frequency.

To model the signal backscattered from a target scene, we discretize the target scene into an $M \times N$ grid in the cross-range and range domains, with each scatterer located at grid (m, n),

Manuscript received May 19, 2015; revised June 24, 2015; accepted June 30, 2015. Date of publication July 02, 2015; date of current version July 09, 2015. This work was supported in part by the National Science Foundation of China under Grants 61201274, 61428103, and by the National Science Foundation under Grant ECCS-1408182. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Glenn Easley.

H. Duan is with the School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: huipingduan@uestc.edu.cn).

L. Zhang, and J. Fang are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China (e-mail: 15196637434@163.com; JunFang@uestc.edu.cn).

L. Huang is with the Department of Electronic and Information Engineering, Shenzhen University, Shenzhen, China (e-mail: dr.lei.huang@ieee.org).

H. Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: Hongbin.Li@stevens.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2015.2452412

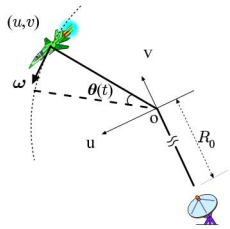

Fig. 1. ISAR model.

and its position and scattering amplitude given respectively by (u_m, v_n) and $x_{m,n}$. In other words, we assume that each range cell consists of M scatterers with different cross-range locations. Hence the signal in the range cell corresponding to $\tau_n = 2(R_0 + v_n)/c$ is a superposition of the echoed signals from the M scatterers [2]

$$y_n(t) = \sum_{m=1}^{M} x_{m,n}\text{rect}(t/T_a)\exp(-j2\pi f_m t) + w_n(t) \quad \forall n \tag{4}$$

where $f_m = 2u_m\omega/\lambda$, $w_n(t)$ denotes the additive noise, and the constant phase term $\exp(-j4\pi(R_0 + v_n)/\lambda)$ is absorbed into the scattering amplitude. By sampling the time series: $t_l = l\Delta t, \forall l = 1, \ldots, L$, where Δt is the pulse repetition period, and $L \triangleq T_a/\Delta t$ is the total number of pulses, the sampled signal $y_n(t_l)$ can be expressed compactly as

$$\boldsymbol{y}_n = \boldsymbol{F}\boldsymbol{x}_n + \boldsymbol{w}_n \quad \forall n \tag{5}$$

where $\boldsymbol{y}_n \triangleq [y_n(t_1) \ \ldots \ y_n(t_L)]^T$, $\boldsymbol{x}_n \triangleq [x_{1,n} \ \ldots \ x_{M,n}]^T$, $\boldsymbol{w}_n \triangleq [w_n(t_1) \ \ldots \ w_n(t_L)]^T$, and $\boldsymbol{F}$ is an $L \times M$ $(L < M)$ matrix with its (l,m)th entry given by $\exp(-j2\pi f_m t_l)$. Putting together signals in different range cells, we obtain

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{X} + \boldsymbol{W} \tag{6}$$

in which $\boldsymbol{Y} \triangleq [\boldsymbol{y}_1 \ \ldots \ \boldsymbol{y}_N]$, $\boldsymbol{X} \triangleq [\boldsymbol{x}_1 \ \ldots \ \boldsymbol{x}_N]$, and $\boldsymbol{W} \triangleq [\boldsymbol{w}_1 \ \ldots \ \boldsymbol{w}_N]$. Clearly, we have $\boldsymbol{Y} \in \mathbb{C}^{L\times N}$, $\boldsymbol{X} \in \mathbb{C}^{M\times N}$, and $\boldsymbol{W} \in \mathbb{C}^{L\times N}$. We assume that $\{\boldsymbol{w}_n\}$ is the additive multivariate Gaussian noise with zero mean and covariance matrix $(1/\gamma)\boldsymbol{I}$. The problem of interest is recover the unknown sparse reflectivity field $\boldsymbol{X}$ from the observed data $\boldsymbol{Y}$.

III. Bayesian Model

To exploit the underlying block-sparse structure in the target image, we propose a two-dimensional pattern-coupled hierarchical Gaussian prior model which is an extension of our previous model [13] to the two-dimensional scenario. Specifically, in the first layer, coefficients $\{x_{m,n}\}$ are assigned a Gaussian prior distribution. To capture pattern dependencies among neighboring coefficients, the Gaussian prior for each coefficient $x_{m,n}$ not only involves its own hyperparameter $\alpha_{m,n}$, but also its immediate neighbor hyperparameters, i.e.

$$p(\boldsymbol{X}|\boldsymbol{\alpha}) = \prod_{m=1}^{M}\prod_{n=1}^{N}\mathcal{CN}(x_{m,n}|0, \delta_{m,n}^{-1}) \tag{7}$$

where $\boldsymbol{\alpha} \in \mathbb{R}^{MN}$ is a hyperparameter vector with its $((n-1)M + m)$th entry equal to $\alpha_{m,n}$, $\mathcal{CN}(\cdot)$ denotes the complex Gaussian distribution, and

$$\delta_{m,n} \triangleq \alpha_{m,n} + \beta \sum_{(i,j)\in N_{(m,n)}} \alpha_{i,j} \tag{8}$$

in which $N_{(m,n)}$ denotes the neighborhood of the grid point (m,n), i.e. $N_{(m,n)} \triangleq \{(m,n-1),(m,n+1),(m-1,n),(m+1,n)\}$[1], and $\beta \in [0,1]$ is a parameter indicating the pattern relevance between the coefficient $x_{m,n}$ and its neighboring coefficients. The second layer specifies Gamma distributions as hyperpriors over the sparsity-controlling hyperparameters

$$p(\boldsymbol{\alpha}) = \prod_{m=1}^{M}\prod_{n=1}^{N}\text{Gamma}(\alpha_{m,n}|a,b) \tag{9}$$

As discussed in [10], for properly chosen a and b, this hyperprior allows the posterior mean of $\alpha_{m,n}$ to become arbitrarily large. As a consequence, the associated coefficients will be driven to zero, thus yielding a sparse solution.

When $\beta = 0$, the prior model (7) reduces to the conventional sparse Bayesian learning model [10]. When $\beta > 0$, we see that the sparsity of each coefficient $x_{m,n}$ is not only controlled by the hyperparameter $\alpha_{m,n}$, but also by the neighboring hyperparameters $\{\alpha_{i,j}|(i,j) \in N_{(m,n)}\}$. The coefficient $x_{m,n}$ will be driven to zero if $\alpha_{m,n}$ or any of its neighboring hyperparameters goes to infinity. Hence the sparsity patterns of neighboring coefficients are related with each other through their shared hyperparameters. On the other hand, for any pair of neighboring coefficients, each of them has its own hyperparameters that are not shared by the other coefficient. It means that no coefficients are pre-specified to share a common sparsity pattern, which enables the prior to provide flexibility to model any block-sparse structures.

Also, the noise variance $1/\gamma$ is assumed unknown, and to estimate this parameter, we place a Gamma hyperprior over γ, i.e.

$$p(\gamma) = \text{Gamma}(\gamma|c,d) \tag{10}$$

where following [10], we set $c = 1$ and $d = 10^{-6}$.

IV. Bayesian Inference

We now proceed to perform Bayesian inference for the proposed pattern-coupled hierarchical model. To facilitate our algorithm development, we first convert (6) as follows by vectorizing the observation matrix $\boldsymbol{Y}$:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w} \tag{11}$$

where $\boldsymbol{y} \triangleq \text{vec}(\boldsymbol{Y})$, $\boldsymbol{x} \triangleq \text{vec}(\boldsymbol{X})$, $\boldsymbol{w} \triangleq \text{vec}(\boldsymbol{W})$, and $\boldsymbol{A} \triangleq \boldsymbol{I} \otimes \boldsymbol{F}$, $\otimes$ stands for the Kronecker product. Clearly, we have $\boldsymbol{y} \in \mathbb{C}^{LN}$, $\boldsymbol{x} \in \mathbb{C}^{MN}$, and $\boldsymbol{A} \in \mathbb{C}^{LN\times MN}$. Based on the above hierarchical model, the posterior distribution of $\boldsymbol{x}$ can be computed as

$$p(\boldsymbol{x}|\boldsymbol{\alpha},\gamma,\boldsymbol{y}) \propto p(\boldsymbol{x}|\boldsymbol{\alpha})p(\boldsymbol{y}|\boldsymbol{x},\gamma) \tag{12}$$

[1]Note that for the edge grid points, they only have two or three immediate neighboring points, in which case the definition of $N_{(m,n)}$ changes accordingly.

According to (7), it can be readily verified that the posterior $p(\boldsymbol{x}|\boldsymbol{\alpha},\gamma,\boldsymbol{y})$ follows a Gaussian distribution with its mean and covariance matrix given respectively by

$$\begin{aligned}\boldsymbol{\mu} &= \gamma\boldsymbol{\Sigma}\boldsymbol{A}^H\boldsymbol{y}\\ \boldsymbol{\Sigma} &= (\gamma\boldsymbol{A}^H\boldsymbol{A}+\boldsymbol{D})^{-1}\end{aligned} \tag{13}$$

where $\boldsymbol{D}$ is a diagonal matrix with its $((n-1)M+m)$th diagonal element equal to $\delta_{m,n}$, i.e.

$$\boldsymbol{D} \triangleq \mathrm{diag}(\delta_{1,1},\ldots,\delta_{M,1},\delta_{1,2},\ldots,\delta_{M,2},\ldots,\delta_{1,N},\ldots,\delta_{M,N})$$

We now exploit the expectation-maximization (EM) formulation to search for the maximum a posterior (MAP) estimate of $\{\alpha_{m,n}\}$ and γ, treating the signal $\boldsymbol{x}$ as hidden variables. Details are elaborated next.

E-Step: Given the current estimates of the hyperparameters $\{\boldsymbol{\alpha}^{(t)},\gamma^{(t)}\}$ and the observed data $\boldsymbol{y}$, the E-step requires computing the expected value (with respect to the hidden variables $\boldsymbol{x}$) of the complete log-posterior of $\{\boldsymbol{\alpha},\gamma\}$, i.e. $E_{\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\alpha}^{(t)},\gamma^{(t)}}[\log p(\boldsymbol{\alpha},\gamma|\boldsymbol{x},\boldsymbol{y})]$, where the operator $E_{\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\alpha}^{(t)},\gamma^{(t)}}[\cdot]$ denotes the expectation with respect to the distribution $p(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\alpha}^{(t)},\gamma^{(t)})$. This complete log-posterior is also referred to as the Q-function. Since

$$p(\boldsymbol{\alpha},\gamma|\boldsymbol{x},\boldsymbol{y}) \propto p(\boldsymbol{\alpha})p(\boldsymbol{x}|\boldsymbol{\alpha})p(\gamma)p(\boldsymbol{y}|\boldsymbol{x},\gamma) \tag{14}$$

the Q-function can be decomposed into a summation of two terms

$$\begin{aligned}Q(\boldsymbol{\alpha},\gamma|\boldsymbol{\alpha}^{(t)},\gamma^{(t)}) &= E_{\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\alpha}^{(t)},\gamma^{(t)}}[\log p(\boldsymbol{\alpha})p(\boldsymbol{x}|\boldsymbol{\alpha})]\\ &\quad + E_{\boldsymbol{x}|\boldsymbol{y},\boldsymbol{\alpha}^{(t)},\gamma^{(t)}}[\log p(\gamma)p(\boldsymbol{y}|\boldsymbol{x},\gamma)]\\ &\triangleq Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(t)},\gamma^{(t)}) + Q(\gamma|\boldsymbol{\alpha}^{(t)},\gamma^{(t)})\end{aligned} \tag{15}$$

Recalling (7) and (13), it can be easily derived

$$\begin{aligned}&Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(t)},\gamma^{(t)})\\ &= \sum_{m=1}^{M}\sum_{n=1}^{N}((a-1)\log\alpha_{m,n} - b\alpha_{m,n} + \log\delta_{m,n} - \delta_{m,n}\omega_{m,n})\end{aligned} \tag{16}$$

where $\omega_{m,n} \triangleq |\hat{\mu}_{m,n}|^2 + \hat{\sigma}_{m,n}$, $\hat{\mu}_{m,n}$ denotes the $((n-1)M+m)$th entry of $\hat{\boldsymbol{\mu}}$, $\hat{\sigma}_{m,n}$ denotes the $((n-1)M+m)$th diagonal element of the covariance matrix $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$ are computed according to (13), with $\boldsymbol{\alpha}$ and γ replaced by the current estimate $\boldsymbol{\alpha}^{(t)}$ and $\gamma^{(t)}$.

The Q-function of γ can be easily obtained as follows

$$Q(\gamma|\boldsymbol{\alpha}^{(t)},\gamma^{(t)}) = (LN+c-1)\log\gamma - \gamma(\eta+d) \tag{17}$$

where

$$\eta \triangleq \|\boldsymbol{y}-\boldsymbol{A}\hat{\boldsymbol{\mu}}\|_2^2 + (\gamma^{(t)})^{-1}\sum_{m=1}^{M}\sum_{n=1}^{N}\rho_{m,n} \tag{18}$$

in which $\rho_{m,n} \triangleq 1-\hat{\sigma}_{m,n}\hat{\delta}_{m,n}$, $\hat{\delta}_{m,n}$ is given by (8), with $\boldsymbol{\alpha}$ replaced by $\boldsymbol{\alpha}^{(t)}$.

M-Step: In the M-step, a new estimate of $\{\boldsymbol{\alpha},\gamma\}$ is obtained by maximizing the Q-function, i.e.

$$\{\boldsymbol{\alpha}^{(t+1)},\gamma^{(t+1)}\} = \arg\max_{\boldsymbol{\alpha},\gamma} Q(\boldsymbol{\alpha},\gamma|\boldsymbol{\alpha}^{(t)},\gamma^{(t)}) \tag{19}$$

We observe that in the Q-function (15), the hyperparameters $\boldsymbol{\alpha}$ and γ are separated from each other. This allows the estimation of $\boldsymbol{\alpha}$ and γ to be decoupled into two independent optimization problems. We first examine the optimization of $\boldsymbol{\alpha}$, i.e.

$$\boldsymbol{\alpha}^{(t+1)} = \arg\max_{\boldsymbol{\alpha}} Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(t)},\gamma^{(t)}) \tag{20}$$

We see that in the Q-function (16), hyperparameters are entangled with each other due to the logarithm term $\log\delta_{m,n}$ (note that $\delta_{m,n}$, defined in (8), is a function of $\boldsymbol{\alpha}$). In this case, an analytical solution to the optimization (20) is difficult to obtain. Gradient descent methods can certainly be used to search for the optimal solution. Nevertheless, here we consider an alternative strategy which aims at finding a simple, analytical sub-optimal solution of (20). Such an analytical sub-optimal solution can be obtained by examining the optimality condition of (20). Suppose $\boldsymbol{\alpha}^* = \{\alpha^*_{m,n}\}$ is the optimal solution of (20), then the first derivative of the Q-function with respect to $\boldsymbol{\alpha}$ equals to zero at the optimal point, i.e.

$$\left.\frac{\partial Q(\boldsymbol{\alpha}|\boldsymbol{\alpha}^{(t)},\gamma^{(t)})}{\partial\alpha_{m,n}}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^*} = \frac{a-1}{\alpha^*_{m,n}} - b + \nu^*_{m,n} - \chi_{m,n} = 0 \tag{21}$$

where

$$\nu^*_{m,n} \triangleq \frac{1}{\delta^*_{m,n}} + \beta\sum_{(i,j)\in N_{(m,n)}}\frac{1}{\delta^*_{i,j}} \tag{22}$$

$$\chi_{m,n} \triangleq \omega_{m,n} + \beta\sum_{(i,j)\in N_{(m,n)}}\omega_{i,j} \tag{23}$$

Since all hyperparameters $\{\alpha_i\}$ and β are non-negative, it can be easily verified $(1/\alpha^*_{m,n}) > (1/\delta^*_{m,n}) > 0$, and $(1/\beta\alpha^*_{m,n}) > (1/\delta^*_{i,j}) > 0$ for $(i,j)\in N_{(m,n)}$. Therefore $\nu^*_{m,n}$ is bounded by $(5/\alpha^*_{m,n}) > \nu^*_{m,n} > 0$. Consequently we have

$$\frac{a+4}{\alpha^*_{m,n}} > \frac{a-1}{\alpha^*_{m,n}} + \nu^*_{m,n} > \frac{a-1}{\alpha^*_{m,n}} \tag{24}$$

Combining (21) and (24), we reach that $\alpha^*_{m,n}$ is within the range

$$\alpha^*_{m,n} \in \left[\frac{a-1}{\chi_{m,n}+b}, \frac{a+4}{\chi_{m,n}+b}\right] \quad \forall m,n \tag{25}$$

A sub-optimal solution to (20) can therefore be simply chosen as

$$\alpha^{(t+1)}_{m,n} = \frac{a-1}{\chi_{m,n}+b} \quad \forall m,n \tag{26}$$

We see that the solution (26) provides a simple rule for the update of $\boldsymbol{\alpha}$. Also, notice that the update rule (26) resembles that of the conventional sparse Bayesian learning work [10] except that $\chi_{m,n}$ is equal to $\omega_{m,n}$ for the conventional sparse Bayesian learning method, while for our case, $\chi_{m,n}$ is a weighted summation of $\omega_{m,n}$ and $\omega_{i,j}$ for $(i,j)\in N_{(m,n)}$. Numerical results show that for properly chosen a and b, this update rule, although sub-optimal, guarantees an exact recovery and provides superior recovery performance.

We now discuss the update of γ. Computing the first derivative of (17) with respect to γ and setting it equal to zero, we obtain

$$\frac{1}{\gamma^{(t+1)}} = \frac{\eta+d}{LN+c-1} \tag{27}$$

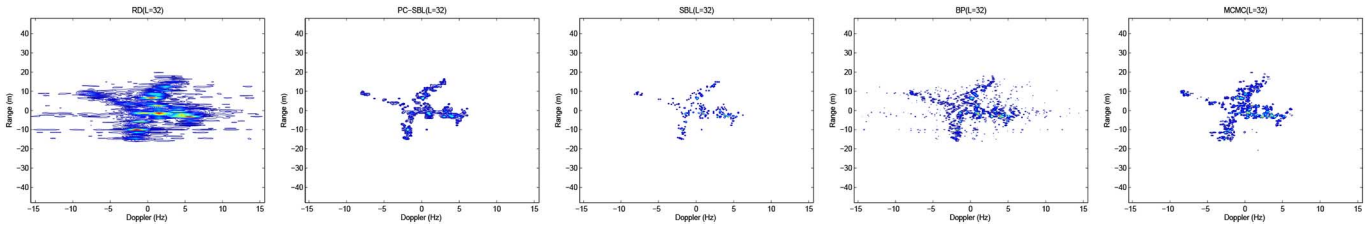

Fig. 2. Images reconstructed from noiseless data by respective algorithms, where $L = 32$.

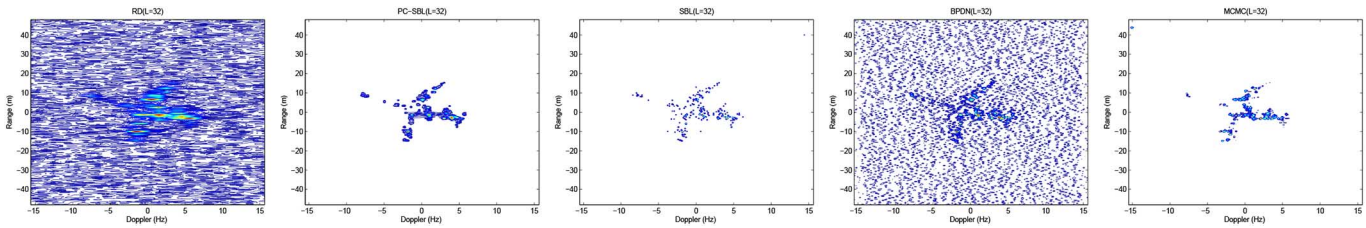

Fig. 3. Images reconstructed from noisy data by respective algorithms, where $L = 32$.

For clarity, we summarize our proposed algorithm as follows.

Algorithm 1 Pattern-Coupled Sparse Bayesian Learning

1. At iteration t ($t = 0, 1, \ldots$): given the current estimates of $\boldsymbol{\alpha}^{(t)}$ and $\gamma^{(t)}$, compute the mean $\hat{\boldsymbol{\mu}}$ and the covariance matrix $\hat{\boldsymbol{\Sigma}}$ of the posterior distribution $p(\boldsymbol{x}|\boldsymbol{\alpha}^{(t)}, \gamma^{(t)}, \boldsymbol{y})$ via (13).
2. Compute a new estimate $\boldsymbol{\alpha}^{(t+1)}$ and $\gamma^{(t+1)}$, according to (26) and (27), respectively;
3. Continue the above iteration until $\|\hat{\boldsymbol{x}}^{(t+1)} - \hat{\boldsymbol{x}}^{(t)}\|_2 \leq \epsilon$, where $\hat{\boldsymbol{x}}^{(t)}$ is chosen as the mean of the posterior distribution $p(\boldsymbol{x}|\boldsymbol{\alpha}^{(t)}, \gamma^{(t)}, \boldsymbol{y})$, and ϵ is a prescribed tolerance value.

V. Simulation Results

We now carry out experiments to illustrate the performance of our proposed pattern-coupled sparse Bayesian learning (PC-SBL) method[2]. We compare our algorithm with the conventional range-Doppler (RD) algorithm [9], the SBL method [10], the basis pursuit (BP/BPDN) method [14], and the MCMC method [6] which exploits the continuity structure of the image. The "Yak-42" plane data set is used in our experiments. The radar system parameters for this data set are the same as that in [7]. The number of range cells is 256, and the total number of pulses is 256 within dwell time of 2.56s. As shown in [6], the underlying reflectivity field of the data set exhibits a block-sparse structure, where strong scatterers are clustered and appear in a plane shape. The parameters a, b and β for our proposed algorithm are set equal to $a = 2$, $b = 10^{-6}$ and $\beta = 1$, where $a = 2$ is empirically proven to be a robust choice for the complex sparse signal case, the choice of β is not very critical to the recovery performance as long as β is set into the range $\beta \in [0.1, 1]$. Also, for PC-SBL and SBL, a pruning operation is adopted, that is, at each iteration, we prune those small coefficients whose associated hyperparameters are greater than $\tau = 10^2$. The corresponding atoms in $\boldsymbol{A}$ are removed from the matrix accordingly. The initial values of $\boldsymbol{\alpha}$ and γ are chosen as $\boldsymbol{\alpha}^{(0)} = 1$ and $\gamma^{(0)} = 10^{-2}$ in our experiments.

[2]Codes are available at http://www.junfang-uestc.net/codes/ISAR.rar

TABLE I
Running Times of Respective Algorithms

Algorithm	PC-SBL	SBL	BP	MCMC
Running Time (sec)	239	229	107	918

We first consider a noiseless case. Fig. 2 shows ISAR images recovered by respective algorithms using $L = 32$ consecutive pulses that are randomly chosen from 256 pulses. We see that when only a small amount of data are available, the conventional RD method provides a defocused and blurred image which is hardly recognized. Also, our proposed PC-SBL method renders the best imaging quality among all methods. It preserves most of the strong scatterers, whereas a considerable portion of scatterers are missing in the image recovered by the SBL method. The BP method yields an obscure image with artifacts outside the target region. We also consider a noisy case where the data are corrupted by an additive white Gaussian noise with noise variance $(1/\gamma) = 9 \times 10^{-4}$(SNR $\approx$ 6 dB). Fig. 3 depicts the images reconstructed by respective algorithms, where $L = 32$. It can be observed that the PC-SBL method still retains decent imaging performance, whereas other compressed sensing methods degrade significantly in the noisy case. The average running times of respective algorithms are also provided in Table I, where results are averaged over 100 independent runs.

VI. Conclusions

We developed a pattern-coupled sparse Bayesian learning method for ISAR imaging. A pattern-coupled hierarchical Gaussian prior was proposed to exploit the block-sparse structure on the target scene. Numerical results show that the proposed method achieves a significant image recovery improvement as compared with existing methods.

REFERENCES

[1] M. Cetin and W. C. Karl, "Feature-enhanced synthetic aperture radar image formation based on nonquadratic regularization," *IEEE Trans. Image Process.*, no. 4, pp. 623–631, Apr. 2001.

[2] L. Zhang, M. Xing, C.-W. Qiu, J. Li, and Z. Bao, "Achieving higher resolution ISAR imaging with limited pulses via compressed sampling," *IEEE Geosci. Remote Sens. Lett.*, no. 3, pp. 567–571, Jul. 2009.

[3] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, no. 6, pp. 2275–2284, Jun. 2009.

[4] V. M. Patel, G. R. Easley, D. M. H. Jr., and R. Chellappa, "Compressed synthetic aperture radar," *IEEE J. Sel. Topics Signal Process.*, no. 2, pp. 244–254, Apr. 2010.

[5] S. Samadi, M. Cetin, and M. A. Masnadi-Shirazi, "Sparse representation-based synthetic aperture radar imaging," *IET Radar, Sonar Navig.*, no. 2, pp. 182–193, 2011.

[6] L. Wang, L. Zhao, G. Bi, C. Wan, and L. Yang, "Enhanced ISAR imaging by exploiting the continuity of the target scene," *IEEE Trans. Geosci. Remote Sens.*, no. 9, pp. 5736–5750, Sep. 2014.

[7] J. Lv, L. Huang, Y. Shi, and X. Fu, "Inverse synthetic aperture radar imaging via modified smooth L_0 norm," *IEEE Antennas Wireless Propogat. Lett.*, pp. 1235–1238, 2014.

[8] M. Cetin, I. Stojanovic, N. O. Onhon, K. R. Varshney, S. Samadi, W. C. Karl, and A. S. Willsky, "Sparsity-driven synthetic aperture radar imaging," *IEEE Signal Process. Mag.*, no. 4, pp. 27–40, Jul. 2014.

[9] J. L. Walker, "Range-Doppler imaging of rotating objects," *IEEE Trans. Aerosp. Electron. Syst.*, no. 1, pp. 23–52, Jan. 1980.

[10] M. Tipping, "Sparse bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, pp. 211–244, 2001.

[11] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.

[12] K.-H. Liu, A. Wiesel, J. David, and C. Munson, "Synthetic aperture radar autofocus based on a bilinear model," *IEEE Trans. Image Process.*, no. 5, pp. 2735–2746, May 2012.

[13] J. Fang, Y. Shen, H. Li, and P. Wang, "Pattern-coupled sparse bayesian learning for recovery of block-sparse signals," *IEEE Trans. Signal Process.*, no. 2, pp. 360–372, Jan. 2015.

[14] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1998.