Research Article | Vol. 25, No. 3 | 6 Feb 2017 | OPTICS EXPRESS 2270

Optics EXPRESS

Compressive sensing and adaptive sampling applied to millimeter wave inverse synthetic aperture imaging

RUOYU ZHU,[1,*] JONATHAN T. RICHARD,[2] DAVID J. BRADY,[1] DANIEL L. MARKS,[1] AND HENRY O. EVERITT[3,4]

[1]*Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA*
[2]*IERUS Technologies, 2904 Westcorp Blvd., Suite 210, Huntsville, AL 35805, USA*
[3]*Department of Physics, Duke University, Durham, NC 27708, USA*
[4]*Army Aviation & Missile RD&E Center, Redstone Arsenal, AL 35898, USA*
[*]*ruoyu.zhu@duke.edu*

Abstract: In order to improve speed and efficiency over traditional scanning methods, a Bayesian compressive sensing algorithm using adaptive spatial sampling is developed for single detector millimeter wave synthetic aperture imaging. The application of this algorithm is compared to random sampling to demonstrate that the adaptive algorithm converges faster for simple targets and generates more reliable reconstructions for complex targets.

OCIS codes: (100.3010) Image reconstruction techniques; (280.6730) Synthetic aperture radar; (280.4750) Optical processing of radar images; (110.6795) Terahertz imaging.

References and links

1. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inf. Theory **52**(2), 489–509 (2006).
2. D. L. L. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**(4), 1289–1306 (2006).
3. X. Yuan, T. Tsai, R. Zhu, P. Llull, D. Brady, and L. Carin, "Compressive hyperspectral imaging with side information," IEEE J. Sel. Top. Signal Process. **9**(6), 964–976 (2015).
4. T. S. Ralston, D. L. Marks, P. S. Carney, and S. A. Boppart, "Interferometric synthetic aperture microscopy," Nat. Phys. **3**(2), 129–134 (2007).
5. D. Takhar, J. N. Laska, M. B. Wakin, M. F. Duarte, D. Baron, S. Sarvotham, K. F. Kelly, and R. G. Baraniuk, "A new compressive imaging camera architecture using optical-domain compression," Proced. Comput. Imaging, SPIE 6065, Computational Imaging **IV**, 606509 (2006).
6. W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," Appl. Phys. Lett. **93**(12), 121105 (2008).
7. L. Xiao, K. Liu, D. Han, and J. Liu, "A compressed sensing approach for enhancing infrared imaging resolution," Opt. Laser Technol. **44**(8), 2354–2360 (2012).
8. O. Furxhi, D. L. Marks, and D. J. Brady, "Echelle crossed grating millimeter wave beam scanner," Opt. Express **22**(13), 16393–16407 (2014).
9. J. Greenberg, K. Krishnamurthy, and D. Brady, "Compressive single-pixel snapshot x-ray diffraction imaging," Opt. Lett. **39**(1), 111–114 (2014).
10. D. J. Brady, D. L. Marks, K. P. MacCabe, and J. A. O'Sullivan, "Coded apertures for x-ray scatter imaging," Appl. Opt. **52**(32), 7745–7754 (2013).
11. D. J. Brady, Optical imaging and spectroscopy. 2009.
12. A. Mrozack, M. Heimbeck, D. L. Marks, J. Richard, H. O. Everitt, and D. J. Brady, "Compressive and adaptive millimeter-wave SAR," Opt. Express **22**(11), 13515–13530 (2014).
13. L. P. Song, C. Yu, and Q. H. Liu, "Through-wall imaging (TWI) by radar: 2-D tomographic results and analyses," IEEE Trans. Geosci. Remote Sens. **43**(12), 2793–2798 (2005).
14. D. M. Sheen, D. L. McMakin, and T. E. Hall, "Three-dimensional millimeter-wave imaging for concealed weapon detection," IEEE Trans. Microw. Theory Tech. **49**(9), 1581–1592 (2001).
15. K. B. Cooper, R. J. Dengler, N. Llombart, T. Bryllert, G. Chattopadhyay, E. Schlecht, J. Gill, C. Lee, A. Skalare, I. Mehdi, and P. H. Siegel, "Penetrating 3-D imaging at 4- and 25-m range using a submillimeter-wave radar," IEEE Trans. Microw. Theory Tech. **56**(12), 2771–2778 (2008).
16. M. S. Heimbeck, D. L. Marks, D. Brady, and H. O. Everitt, "Terahertz interferometric synthetic aperture tomography for confocal imaging systems," Opt. Lett. **37**(8), 1316–1318 (2012).
17. E. Yiğit, "Compressed sensing for millimeter-wave ground based SAR/ISAR imaging," J. Infrared Millim. Terahertz Waves **35**(11), 932–948 (2014).

#278540 https://doi.org/10.1364/OE.25.002270
Journal © 2017 Received 12 Oct 2016; revised 10 Jan 2017; accepted 10 Jan 2017; published 27 Jan 2017

Optics EXPRESS

18. W. L. Chan, J. Deibel, and D. M. Mittleman, "Imaging with terahertz radiation," Rep. Prog. Phys. **70**(8), 1325–1379 (2007).
19. M. Martorella, J. Palmer, F. Berizzi, and B. Bates, "Advances in Bistatic Inverse Synthetic Aperture Radar," Radar Conference - Surveillance for a Safer World, 2009. RADAR. International. pp. 1–6, 2009.
20. R. Baraniuk and P. Steeghs, "Compressive radar imaging," in IEEE National Radar Conference Proceedings, 128–133 (2007)
21. L. Yu and Y. Zhang, "Random step frequency CSAR imaging based on compressive sensing," Prog. Electromagn. Res. C **32**, 81–94 (2012).
22. J. H. G. Ender, "On compressive sensing applied to radar," Signal Process. **90**(5), 1402–1414 (2010).
23. S. Ji, Y. Xue, and L. Carin, "Bayesian Compressive Sensing," IEEE Trans. Signal Process. **56**(6), 2346–2356 (2008).
24. M. Tipping, "Sparse Bayesian Learning and the Relevance Vector Machine," J. Mach. Learn. Res. **1**, 211–214 (2001).
25. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process. **13**(4), 600–612 (2004).
26. I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," IEEE Trans. Inf. Theory **36**(5), 961–1005 (1990).
27. G. Lauritsch and W. H. Haerer, "Theoretical framework for filtered back projection in tomosynthesis," Proc. SPIE **3338**, 1127–1137 (1998).

1. Introduction

Compressive sensing (CS) algorithms [1, 2] are increasingly being adapted for image acquisition because they may increase the dimension of images, such as by adding side information to resolve spatial images spectrally and temporally as well [3]. They also significantly improve single detector imaging methods by increasing the sampling speed and efficiency over raster-scanning [20], as demonstrated in applications as diverse as optical coherence tomography (OCT) [4], single photon imaging [5, 6], infrared imaging [7], millimeter wave imaging [8], and x-ray imaging [9, 10]. CS achieves this by overcoming the Nyquist sampling limit, requiring only a few samples to reconstruct the original signal if it is sparse on some basis, such as the wavelet basis [26]. Instead of measuring a full data set, the single pixel detector may only need to acquire a few well-chosen samples to achieve a reliable reconstruction. Consequently, sampling strategies [11] have become a topic of great interest, especially CS-based adaptive sampling algorithms [12]. These are sequential sampling methods that improve the sampling efficiency by inferring which points to sample next using knowledge of prior samples.

Growing in popularity as the technology matures, millimeter wave (MMW) imaging techniques are increasingly attractive for non-destructive high-resolution imaging and radar applications such as through the wall imaging [13], concealed weapon detection [8, 14, 15], and foreign object detection [16, 17]. High-resolution MMW imaging requires broad bandwidths [18] and large apertures [14]. Because of the unavailability of sensitive, low-cost detector arrays, MMW imaging relies heavily on synthesizing large apertures to acquire high-resolution images. Synthetic aperture radar (SAR) and inverse synthetic aperture radar (ISAR) images are acquired by mechanically moving a single aperture [16, 19] at the cost of slow acquisition speeds. Specifically, SAR images are formed by moving an aperture along a straight path and illuminating a stationary object, while ISAR images are either acquired by moving the aperture along a circular path around the target or by rotating the object in front of a fixed aperture. Although synthesizing a large aperture by mechanically moving a single pixel detector reduces the cost for a physically large aperture, mechanical scanning requires a long acquisition time, and mechanical stability becomes an issue as large volumes of spatial-spectral data are acquired.

CS algorithms have already been applied to MMW SAR and ISAR imaging [17, 20–22]. It was shown in [17, 20] that compression can be done on the spatial and spectral data by random sampling and reconstructions using various CS algorithms. The primary random sampling strategies include random spatial sampling, random spectral sampling, and random spatial-spectral sampling. Random spatial-spectral sampling provides the best compression rate because of the huge amount of spatial-spectral data acquired. However, if the objective is

Optics EXPRESS

to minimize acquisition time by reducing the amount of mechanical scanning, random spatial sampling is far superior to random spectral and random spatial-spectral sampling, even though the data compression rate is lower.

When MMW imaging is being used to image through obscurants, the complexity of the scene is unknown [12–15]. Consequently, random sampling strategies are problematic because the number of samples needed cannot be predetermined. Adaptive sampling provides a compelling alternative that optimizes the measurement scheme regardless of the scene complexity, thereby minimizing the amount of spatial scanning required, especially for ISAR. In a proof-of-concept demonstration of an adaptive sampling CS algorithm applied to MMW SAR imaging, Mrozack et. al accurately located point scatterers in as many steps, but the algorithm was of limited utility because of its requirement that the exact number of scatterers in the scene must be known in advance [12]. Here we introduce an ISAR method that adaptively selects each measurement location based on the Bayesian compressive sensing (BCS) framework [23, 24] and needs no prior information on the scene complexity. The need for BCS applied to MMW ISAR was borne out of a need for a faster way to obtain high quality reconstructions of complex targets using a single heterodyne MMW transceiver. The current methodologies require mechanically scanning the target sequentially through many angles. This mechanical scanning is slow and further wastes time by measuring many angles that provide little critical information for the reconstruction. In contrast, range is determined by rapid electronic frequency sweeps, so there is no need for BCS in the range dimension. Therefore, by developing a methodology to identify which few angles can provide the most critical information needed for a reconstruction, we can avoid measuring angles that provide little additional information and speed up acquisition. We show that the adaptive algorithm converges faster than random sampling for simple targets and generates more reliable reconstructions for complex targets. In addition, the BCS framework allows the user to define stopping criteria without prior knowledge of the scene.

2. Method

Other 2D imaging modalities directly sample in the x-y domain to render a target scene with reflectivity $\rho(x,y)$, but ISAR (and computational tomography (CT)) are sampled in the angle-range domain $f(\theta,\mathrm{r})$. Figure 1(a) illustrates how this is done for traditional MMW ISAR imaging in which either the target or the imaging system is rotated and the interrogating beam from a transceiver horn antenna is directed toward the center of the scene at a sequence of angles θ_i and frequencies ν. As shown in Fig. 1(b), the same transceiver antenna measures the reflected signal $F(\theta_i,\nu)$ using a vector network analyzer to record the S11 parameter at each angle over a range of frequencies. The transceiver antenna is then moved to the next angle θ_{i+1} and the measurement process is repeated until $F(\theta_i,\nu)$ is acquired at all angles and frequencies. From this we can represent the scene in the desired $f(\theta,\mathrm{r})$ angle-range domain, where range r is measured from the center of the scene by a simple Fourier transform

$$\mathrm{f}(\theta,\mathrm{r}) = \int F(\theta,\nu) e^{j2\pi\nu r} d\nu. \tag{1}$$

Figure 1(c) plots an example of $f(\theta,\mathrm{r})$, which bears a resemblance to the sinograms acquired in CT imaging as scattering elements in the scene rotate closer or farther from the transceiver. Since the range data represents the strength of the reflected signal in a given direction, $f(\theta,\mathrm{r})$ is related to $\rho(x,y)$ by the Radon transform

$$f(\theta,r) = \int\int \rho(x,y)\delta(x\cos\theta + y\sin\theta - r)dxdy. \tag{2}$$

Optics EXPRESS

The reflectivity function can be easily reconstructed by a filtered back projection (FBP) [27]

$$\rho(x,y)=\int_0^{\pi} f'(\theta, x\cos\theta+y\sin\theta)d\theta. \tag{3}$$

where $f'(\theta,r)$ denotes the range data that has been high-pass filtered with a ramp filter to improve the result. Figure 1(d) illustrates the reconstruction obtained by applying the FBP algorithm to the range data in Fig. 1(c).

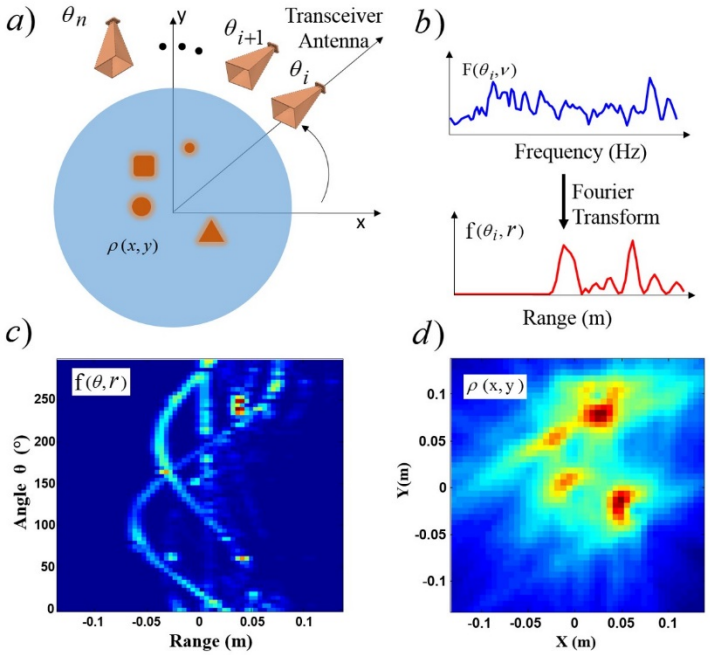

Fig. 1. a) Imaging configuration for traditional ISAR imaging. b) Example of a measurement at a single angle in the frequency domain $F(\theta_i,\nu)$ and the range data in the spatial domain obtained by Fourier transform. c) Complete range data of all angles $f(\theta,\mathrm{r})$. d) Reconstruction of the reflectivity function $\rho(x,y)$ by filtered back projection.

To model the measurement process, we describe the forward model as

$$g(\theta,\mathrm{r})=\iint h(\theta-\theta',\mathrm{r}-\mathrm{r}')\,\mathrm{f}(\theta',r')d\theta'dr'. \tag{4}$$

where $g(\theta,\mathrm{r})$ is the subset of $f(\theta,\mathrm{r})$ sampled through the transfer function $h(\theta,r)$ that describes the sampling scheme. The model should be discretized since the samples are discretized, so

$$g_\Delta(\theta,r)=\sum_{n=1}^{N}\sum_{m=1}^{M}h_\Delta(\theta-m\Delta\theta',\mathrm{r}-n\Delta r')\,\mathrm{f}(\mathrm{m}\,\Delta\theta',n\Delta r'). \tag{5}$$

where $\Delta\theta$ and Δr denotes the sampling interval in the angular and range domain, respectively, $M=2\pi/\Delta\theta$ is the total number of angular samples, and $N=R/\Delta r$ is the total number of range samples. From Eq. (5) a vectorized expression of the model becomes

$$\mathbf{g}=\mathbf{H}\mathbf{f}. \tag{6}$$

Optics EXPRESS

where $\mathbf{g}$ is the measurement vector of size MNx1, $\mathbf{H}$ is the MN x MN forward matrix, and $\mathbf{f}$ is the MN x1 vector as the target for reconstruction. For the ISAR scenario illustrated in Fig. 1(a), the transceiver measures range data by means of a frequency sweep at angle θ_m, so if a single element in $\mathbf{g}$ is denoted $g(\theta,r)\big|_{\substack{\theta=m\Delta\theta\\ r=n\Delta r}}$ and labeled $\mathbf{g_{mn}}$, the data collected at angle θ_m become $\mathbf{g_m}=[\mathbf{g_{m1}};\mathbf{g_{m2}};...;\mathbf{g_{mn}}]$. The corresponding transfer function may be described as $\mathbf{h_m}=[\mathbf{h_{m1}},\mathbf{h_{m2}},...\mathbf{h_{mn}}]^T$, where $\mathbf{h_{mn}}=\delta(\theta-m\Delta\theta,r-n\Delta r)$ describes the sampling of a point in the angle-range domain. The objective is to minimize the time required to obtain $\mathbf{f}$.

To see how this is done, we next discuss the adaptive sampling strategy under the compressive sensing framework. Instead of sampling at all M angles, $\mathbf{f}$ can be sampled incompletely with only K samples (K<M) and be reconstructed using the compressive sensing algorithm. We assume $\mathbf{f}$ needs to be sparse or sparse under wavelet representation for compressive reconstruction. Thus we chose to represent $\mathbf{f}$ by the Haar wavelets, resulting in the wavelet coefficients $\boldsymbol{\omega}$. We denote the Haar wavelet basis as an $MN\times MN$ matrix $\mathbf{B}$, so that $\mathbf{f}=\mathbf{B}\boldsymbol{\omega}$. The compressive measurement can thus be expressed as $\mathbf{g}=\mathbf{HB}\boldsymbol{\omega}=\boldsymbol{\Phi}\boldsymbol{\omega}$, where $\boldsymbol{\Phi}=\mathbf{HB}=[\mathbf{r_1}...\mathbf{r_K}]^{\mathrm{T}}$ is the projection matrix relating the wavelet coefficients $\boldsymbol{\omega}$ to the compressive measurements, and a single measurement is simply $\mathbf{g_k}=\mathbf{h_k}^{\mathrm{T}}\mathbf{B}\boldsymbol{\omega}=\mathbf{r_k}^T\boldsymbol{\omega}$. Note that here we use the notation $\mathbf{g_k}$ instead of $\mathbf{g_m}$ to differentiate that the former corresponds to the compressive measurement. The sparse measurement constitutes an ill-posed inversion problem, which is commonly solved via L-1 regularization, i.e.

$$\hat{\omega}=\arg\min_{\omega}\left\{\left\|\mathbf{g}-\Phi\omega\right\|_2^2+\lambda\left\|\omega\right\|_1\right\}. \tag{7}$$

where λ is the regularization coefficient that controls the sparseness of the estimation result. However, the L-1 regularization only provides a point estimate, which gives no access to adaptively choosing the optimal next measurement. Others [23, 24] have formulated the CS inversion from a Bayesian learning point of view, providing a posterior density function on the signal under reconstruction. Besides the improved accuracy of point estimation, the full posterior density function results in confidence intervals on the estimation and thus provides an objective to optimize the next measurement. Therefore the BCS framework is not only able to reconstruct the original data compressively using significantly fewer measurements, it also provides a way to find the next projection $\mathbf{r_{k+1}}$ after each measurement. This property of the BCS framework makes it a popular choice to achieve adaptive sampling in many imaging modalities. This paper focuses on how to use the full posterior density function to form the next projection $\mathbf{r_{k+1}}$ in the projection matrix Φ, then how to realize the corresponding sampling scheme $\mathbf{h_{k+1}}$ in ISAR. The theoretical framework for the BCS algorithm is next outlined, following [23, 24], to explain how it reconstructs the original signal and formulates the optimization method for choosing the next measurement angle.

The compressive measurement can be expressed as the projection of the significant wavelet coefficients ω_s, the remaining insignificant wavelet coefficients, ω_e, and the measurement noise $\mathbf{n_m}$. Thus the measurement $\mathbf{g}$ can be expressed as

$$\begin{aligned}\mathbf{g}&=\Phi\omega+\mathbf{n}_m\\&=\Phi\omega_s+\Phi\omega_e+\mathbf{n}_m.\\&=\Phi\omega_s+\mathbf{n}\end{aligned} \tag{8}$$

The elements in the generalized measurement noise $\mathbf{n}$ are approximated as zero-mean Gaussian distribution with variance σ^2. Thus, the likelihood for the measurement is

Optics EXPRESS

$$p(\mathbf{g} \mid \omega_s, \sigma^2) = (2\pi\sigma^2)^{-K/2} \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{g} - \Phi\omega_s\|^2 \right). \tag{9}$$

For a compressive sensing problem where the signal is assumed to be sparse in the wavelet basis, a sparseness prior must be placed in the Bayesian formulation. A widely used sparseness prior is the Laplace density function [21, 23, 24]. The conventional CS conversion resulting in the solution of Eq. (7) can be seen as the maximum *a posteriori* (MAP) estimation of $\boldsymbol{\omega}$ [23, 24]. However, using the Laplace density function as a sparsity prior results in a Bayesian inference that cannot be calculated in closed form because the Laplace prior is not conjugate to the Gaussian likelihood shown in Eq. (9). Previous authors [23, 24] have provided a solution using the relevance vector machine (RVM) framework that imposes a hierarchical prior with similar properties but are conjugates to the Gaussian likelihood in Eq. (9). A zero-mean Gaussian prior is first defined on ω,

$$p(\omega \mid \boldsymbol{\alpha}) = \prod_i^N \mathcal{N}\left(\omega_i \mid 0, \alpha_i^{-1}\right). \tag{10}$$

where α_i is the inverse variance indicating the precision. The second level is a Gamma prior on the hyper-parameter $\boldsymbol{\alpha}$,

$$p(\boldsymbol{\alpha} \mid \mathrm{a}, \mathrm{b}) = \prod_i^N \Gamma\left(\alpha_i \mid \mathrm{a}, \mathrm{b}\right). \tag{11}$$

The final prior on ω is thus derived by marginalizing over $\boldsymbol{\alpha}$,

$$p(\omega \mid a, b) = \prod_i^N \int_0^\infty \mathcal{N}\left(\omega_i \mid 0, \alpha_i^{-1}\right) \Gamma\left(\alpha_i \mid \mathrm{a}, \mathrm{b}\right) d\alpha_i. \tag{12}$$

Note that the integral in Eq. (12) results in the Student-t distribution [24]. This prior function can thus promote sparseness by choosing the right values for **a** and **b** so that it reaches maximum when $\omega_i = 0$.

For the noise in the measurement, i.e. **n** in Eq. (8), a similar hierarchical prior is placed with α_0 as the inverse variance of the noise,

$$p(\mathbf{n} \mid a, b) = \prod_i^K \int_0^\infty \mathcal{N}\left(n_k \mid 0, \alpha_0^{-1}\right) \Gamma\left(\alpha_0 \mid \mathrm{c}, \mathrm{d}\right) d\alpha_i. \tag{13}$$

The posterior covariance and mean are respectively given by [15]

$$\begin{aligned} \boldsymbol{\Sigma} &= (\alpha_0 \Phi^T \Phi + A)^{-1} \\ \boldsymbol{\mu} &= \alpha_0 \Sigma \Phi^T \mathbf{g} \end{aligned} \tag{14}$$

where $\mathbf{A} = diag(\alpha_0, \alpha_1, ..., \alpha_N)$. The goal thus becomes searching for the precision $\boldsymbol{\alpha}$ and α_0. The marginal likelihood for $\boldsymbol{\alpha}$ and α_0 can be derived by marginalizing over the weights $\boldsymbol{\omega}$, following a type-II maximum likelihood (ML) procedure [24], i.e.,

$$\begin{aligned} \mathcal{L}(\boldsymbol{\alpha}, \alpha_0) &= \log p(\mathbf{g} \mid \boldsymbol{\alpha}, \alpha_0) \\ &= \log \int p(\mathbf{g} \mid \boldsymbol{\omega}, \alpha_0) p(\boldsymbol{\omega} \mid \boldsymbol{\alpha}) \, \mathrm{d}\boldsymbol{\omega}. \\ &= -\frac{1}{2}[\mathrm{K} \log 2\pi + \log |\mathbf{C}|] + \mathbf{g}^T \mathbf{C}^{-1} \mathbf{g} \end{aligned} \tag{15}$$

Optics EXPRESS

where $\mathbf{C} = \sigma^2\mathbf{I} + \mathbf{\Phi}\mathbf{A}^{-1}\mathbf{\Phi}^T$. The point estimates for $\boldsymbol{\alpha}$ and α_0 from type II-ML maximizing the likelihood yields

$$\begin{aligned} \alpha_i^{new} &= \frac{\gamma_i}{\mu_i^2} \\ \alpha_0^{new} &= \frac{K - \Sigma_i \gamma_i}{\| \mathbf{g} - \mathbf{\Phi}\boldsymbol{\mu} \|_2^2}. \end{aligned} \tag{16}$$

where μ_i is the i-th posterior mean from Eq. (8), and $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ [23, 24]. $\boldsymbol{\alpha}$ and α_0 are then used to calculate $\boldsymbol{\Sigma}$ and $\boldsymbol{\mu}$ using Eq. (14) again, which constitutes an iterative process that updates these parameters until convergence occurs. The fact that $\boldsymbol{\alpha}$ and α_0 can be updated iteratively indicates that initializing the **a, b, c, d** parameters for the hierarchical Gamma priors in Eq. (12) and Eq. (13) is not necessary in this case. Thus, setting them equal to zero is equivalent to enforcing a uniform prior on $\boldsymbol{\alpha}$ and α_0 [23].

Since the signals being reconstructed are the wavelet coefficient of the original signals, where $\mathbf{f} = \mathbf{B}\omega$, the expectation and covariance of the posterior density function can respectively be derived as

$$\begin{aligned} E(\mathbf{f}) &= \mathbf{B}\boldsymbol{\mu} \\ Cov(\mathbf{f}) &= \mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^{\mathbf{T}} \end{aligned} \tag{17}$$

where the diagonals of $\boldsymbol{\Sigma}$ indicate the level of accuracy (or uncertainty) on the reconstruction of elements in $\mathbf{f}$ [23].

The ability to measure uncertainty provides criteria for selecting the next best measurement to minimize the total number of samples needed to reconstruct the signal with significant fidelity. The differential entropy is one criterion that satisfies the purpose [23, 24], where

$$\begin{aligned} \mathrm{S(f)} &= -\int p(f)\log p(f)df \\ &= \frac{1}{2}\log|\mathbf{B}\boldsymbol{\Sigma}\mathbf{B}^T| + \mathrm{const}. \\ &= \frac{1}{2}\log|\boldsymbol{\Sigma}| + \mathrm{const} \\ &= -\frac{1}{2}\log|\mathbf{A} + \alpha_0\mathbf{\Phi}^{\mathbf{T}}\mathbf{\Phi}| + \mathrm{const} \end{aligned} \tag{18}$$

With a new measurement, $\mathbf{\Phi}$ is modified by adding a new row r_{k+1}^T; thus, the new entropy after the next measurement derived in [23] is

$$S_{new}(\mathrm{f}) = h(\mathrm{f}) - \frac{1}{2}\log|\mathbf{A} + \alpha_0\mathbf{r}_{K+1}{}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}|. \tag{19}$$

The goal of making the next possible measurement becomes defining $\mathbf{r}_{K+1}$ so that the new entropy $S_{new}(\mathrm{f})$ is minimized, so the maximum of $\mathbf{r}_{K+1}{}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$ should be pursued. In addition, as pointed out in [14], maximizing $\mathbf{r}_{K+1}{}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$ is equivalent to maximizing the variance of the next measurement, since

$$\begin{aligned} \mathbf{r}_{K+1}{}^{\mathbf{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1} &= \mathbf{r}_{K+1}{}^{\mathbf{T}}Cov(\omega)\mathbf{r}_{K+1}. \\ &= Var(g_{K+1}) \end{aligned} \tag{20}$$

Optics EXPRESS

The user may specify a stopping criterion for the algorithm to end based on the desired amount that $\mathbf{r}_{K+1}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$ changes from iteration to iteration.

Now the question becomes how to find the $\mathbf{r}_{K+1}$ that maximizes the variance $\mathbf{r}_{K+1}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$. Theoretically, the next projection vector $\mathbf{r}_{K+1}$ should be designed by performing an eigen-decomposition of $\boldsymbol{\Sigma}$ and letting $\mathbf{r}_{K+1}$ be the eigenvector of the largest eigenvalue [23]. Recall that the measurement and target are related by $\mathbf{g_k} = \mathbf{h_k}^{\mathrm{T}}\mathbf{f} = \mathbf{h_k}^{\mathrm{T}}\mathbf{B}\boldsymbol{\omega} = \mathbf{r_k}^{T}\boldsymbol{\omega}$, so to measure the target in its original non-wavelet basis, we use the relationship $\mathbf{h_{k+1}} = \mathbf{B}^{T}\mathbf{r_{k+1}}$ to find the $\mathbf{h_{k+1}}$ and design the sampling method. Here, $\mathbf{h_{k+1}}$ is in the form of multiplexed sampling given by

$$\mathbf{h}_{k+1} = \sum_i \sum_j \delta(\theta - i\Delta\theta, \mathrm{r} - j\Delta r). \tag{21}$$

An example illustrating how the projection vector $\mathbf{h_{k+1}}$ is realized under a hypothetical traditional BCS framework is presented in Fig. 2(a). The colored points in Fig. 2(a) indicate which points are being sampled in each measurement: orange points depict the transfer function of the k-th measurement $h_k(\theta, r)$ as an overlay on the target function $f(\theta, r)$, and the green points depict the next measurement $h_{k+1}(\theta, r)$. The formation of the forward matrix $\mathbf{H}$ is thus visualized in Fig. 2(c) where each row corresponds to a measurement, and multiple points in each row indicate multiplexing. This type of sampling is usually realized with a coded aperture [11] to modulate which pixels to sample. However, this sampling strategy is not practical for ISAR imaging because the transceiver can only measure from one angle at a time. The sampling described in Eq. (21) requires measuring multiple angles and ranges at the same time, requiring the kind of extensive mechanical movement that we are trying to avoid.

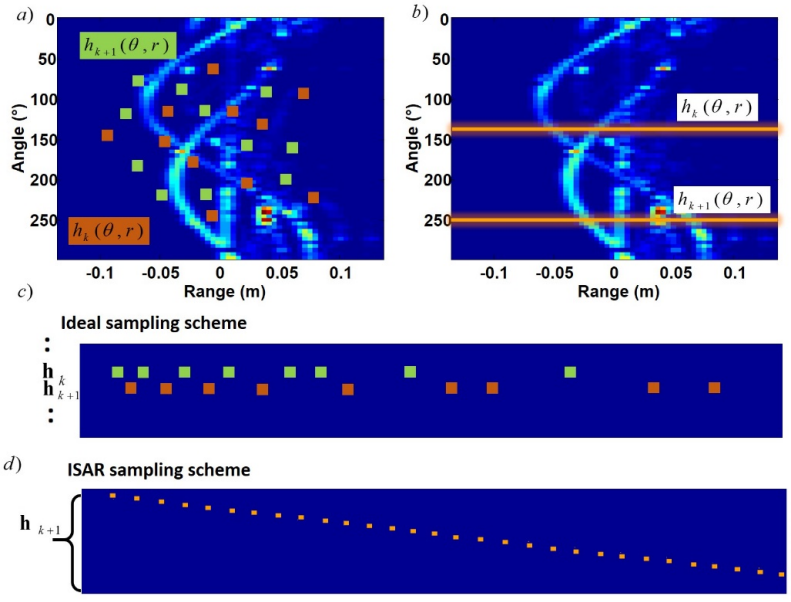

Fig. 2. a) Visualization of the next measurement under a hypothetical traditional BCS methodology. Colored points indicate the points sampled in step k (orange) and k + 1 (green). b) Visualization of the next measurement for ISAR imaging. c) Vectorized form of $\mathbf{h}_{K+1}$ for ideal BCS, and d) $\mathbf{h}_{K+1}$ for ISAR imaging.

Optics EXPRESS

As noted above, ISAR imposes a constraint when maximizing $\mathbf{r}_{K+1}{}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{r}_{K+1}$: all range data are taken for a fixed angle during each measurement, and this is illustrated in Fig. 2(b). Given that our goal is to minimize the amount of mechanical movement, the sampling choice is thus limited to selecting the next angle θ_i. Due to this constraint, the actual next projection for ISAR imaging cannot be the exact eigenvector that enforces the maximum of $\mathbf{r}_{K+1}{}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{r}_{K+1}$, and we lose the advantage of multiplexing the measurement. As a result, the transfer function should be described as

$$\mathbf{h}_{k+1}^{ISAR} = [\mathbf{h}_{k+1,1}, \mathbf{h}_{k+1,2}, \ldots, \mathbf{h}_{k+1,\mathrm{n}}]^T \tag{22}$$

$$\mathbf{h}_{k+1,\mathrm{n}} = \delta(\theta - \theta_{k+1}, \mathrm{r} - n\Delta r). \tag{23}$$

Figure 2(b) and 2(d) illustrates the sampling strategy by visualizing the transfer function for the ISAR measurement and its corresponding vectorized form, the transfer matrix. The difference between the typical BCS framework and the application of BCS to ISAR is that there is no multiplexing in the latter [11], so more measurements are required. Therefore, to select the next optimal measurement in ISAR, we have to define a library of all possible $\mathbf{h}_{K+1}$ matrices under to the predefined sampling intervals and choose the one that maximizes the sum of $\mathbf{r}_{K+1}{}^{\mathrm{T}}\mathbf{\Sigma}\mathbf{r}_{K+1}$. For example if the user defines 360 samples along the angular domain, such that the angular sampling interval is 1° and M = 360, the library consists of 360 possible $\mathbf{r}_{K+1}$ matrices from the beginning. Given the $\mathbf{\Sigma}$ from current measurements, the next measurement is chosen from the remaining 359 angles. Figures 2(c) and 2(d) emphase the difference between the theoretical BCS framework and the ISAR BCS framework that is constrained from multiplexing. Under a hypothetical BCS framework with complete freedom to multiplex, the size of the H matrix with K measurements is K x MN, and since K can be a very small number, the compression rate may be maximized. However, for the ISAR application the BCS framework has a limited ability to multiplex and can only adaptively choose among K angles, so the size of the H matrix becomes a much larger KN x MN.

3. Experiments and results

Two sets of experiments were performed to test the adaptive sensing algorithm. For both experiments, the sample objects are mounted on a motorized rotational stage. The objects are illuminated by a 15 cm diameter beam folded by and collimated with a mirror. The beam is reflected by the sample object and measured by a single stationary detector. Figure 3(a) presents a schematic drawing of the experiment setup. The source radiation is generated by a transceiver module frequency swept from 75 to 110 GHz. The reflected signal is analyzed with a network analyzer. The spatial measurement is done by rotating the target on the rotational stage, then obtaining the spectral data and evaluating its Fourier transform.

Optics EXPRESS

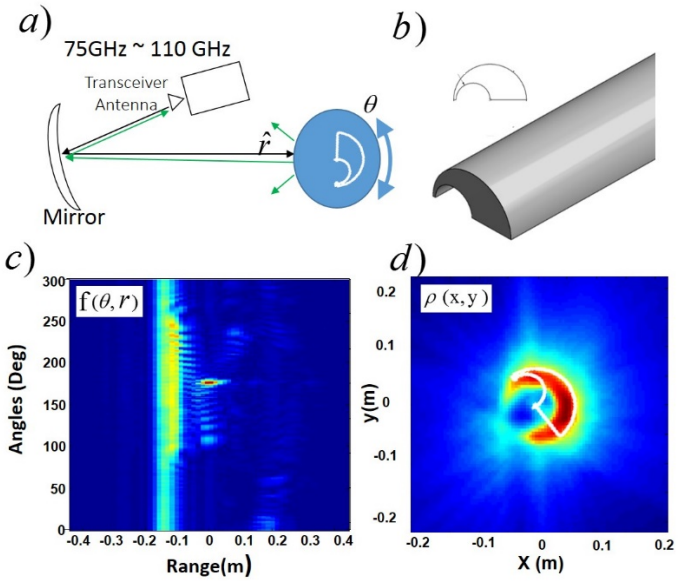

Fig. 3. (a) A schematic drawing of the experiment. (b) Drawing of the target half-cylinder nautilus. (c) The range signal measured at 300 angles separated at 0.1 degrees per measurement. The fully-sampled reconstruction is treated as the ground truth for the simulated experiments. (d) The ISAR reconstruction via the filtered backpropagation method with all 3000 measurements.

The sample under investigation is a 3D printed nautilus-shaped dielectric cylinder, shown in Fig. 3(b). The cylinder is 20 cm tall with outer radius 30 mm and inner radius 14 mm. Figure 3(c) plots the spatial-spectral (i.e. angle-range) data measured at 3000 locations equally spaced between from 0o to 300o. The 180° orientation is depicted in Fig. 3(a), and the strong specular reflection from the flat surface is easily seen in Fig. 3(c). The range-varying structure between 90°-270° is produced by reflections from the inner cylinder cavity as it rotates closer to the source. The ISAR image can be reconstructed by performing filtered back-projection using the inverse Radon transform on the measured spatial data. The results, shown in Fig. 3(d), indicate a faithful reconstruction of the outer half-cylinder surface but a less accurate reconstruction of the "eclipsed" flat surface because of multiple reflection. The full sampling measurements and reconstruction shown in Figs. 3(c) and 3(d) constitute the reference to which the adaptively sampled and reconstructed images will be compared.

The adaptive sensing algorithm described here is based on a Bayesian compressive sensing framework [23, 24], which requires the sparsity of the signal. However, the target under investigation may not satisfy the sparsity requirement. We perform a wavelet transform (Haar) on the measured signal in the range domain, which is relatively sparse compared to the signal in the spatial domain, shown in Fig. 4(a). The wavelet transform provides a representation of the signal in a multi-scale basis [26]. For wavelets with lower order scales, the original signal is represented by a slowly varying basis that produces a few coefficients with large values. For higher order wavelets, the original signal is represented by a version of the original basis scaled with faster variation and smaller windows, generating a large number of coefficients with much smaller values. In Fig. 4(a), the wavelet coefficients for the target are shown in a vectorized form. The lower order coefficients correspond to the part of the signal with slower variations, whereas the higher order coefficients correspond to faster variations and finer features. The spectrum of the wavelet coefficient also implies the sparsity of the target.

Optics EXPRESS

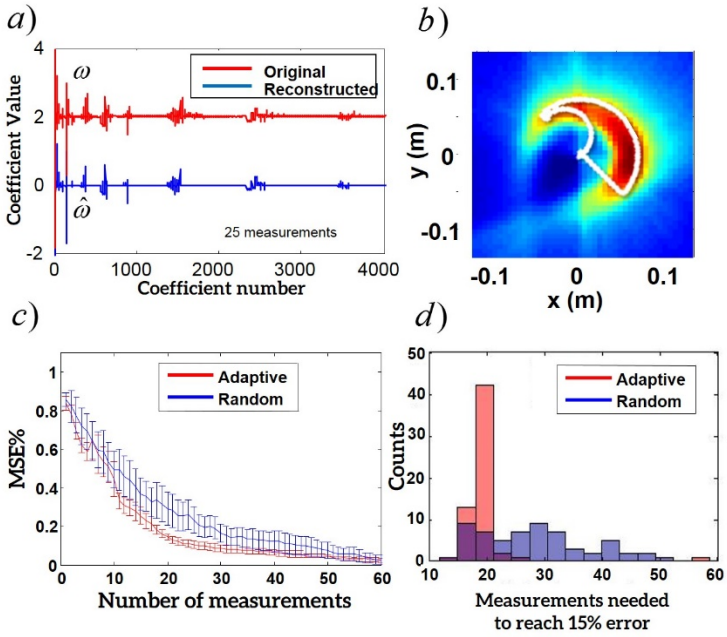

Fig. 4. (a) The reconstruction of the wavelet coefficient of the signal after 25 measurements. (b) The ISAR reconstruction via filtered back projection on the data with 25 measurements. (c) The average MSE comparison of adaptive spatial sampling and random spatial sampling for 60 simulated experiments. Error bars indicate the standard deviation of the MSE. (d) Histogram of how many measurements the adaptive and random sampling has taken to reach 15% MSE.

The reconstructed signal after 25 adaptive measurements is shown as an offset to the original signal in Fig. 4(a). From the comparison between the wavelet coefficients of the full measurement and that of the compressed measurement, it can be seen in Fig. 4(a) that the reconstruction with the first 25 measurements selected by the adaptive algorithm can restore the original signal. The minor difference in higher order coefficients can be tolerated because they are likely to be the result of noise in the measurement and contribute only a small percentage in terms of energy. The first 25 measurements in the range domain are reconstructed and back-projected to form the target image shown in Fig. 4(b). The white outline depicts the true contour of the target. To assess the reconstruction quality of the compressed adaptive measurement, we measured a 0.944 structural similarity index (SSIM) [25] when comparing to the full measurement reconstruction $\rho(x,y)$ in Fig. 3(d). It should be pointed out that the BCS algorithm allows the user to specify a stopping criteria when the change of uncertainty in the measurement is not increasing any more, yet we let the simulated experiment run without stopping to analyze convergence. The goal here is to provide information that can be used to define appropriate stopping criteria.

The idea of using CS techniques to reconstruct the target image from a small subset of all possible measurements has already been demonstrated [1, 2]. It was shown that randomly selecting measurement angles in an ISAR experiment might substantially reduce the number of measurements and thus reduce the amount of mechanical scanning. Using an adaptive algorithm for compressed spatial sampling further improves the efficiency by achieving an acceptable reconstruction faster.

To demonstrate this advantage, we have simulated the experiment 60 times, each starting at a different angle, and compared it to the results of 60 different experiments using random measurements. The result is shown in Fig. 4(c), using the normalized mean squared error

Optics EXPRESS

(MSE) $\frac{1}{MN}\sum_{i=1}^{MN}(\mathbf{f_i}-\hat{\mathbf{f}}_\mathbf{i})^2$ normalized by $\|\mathbf{f}\|^2$ as the reference criterion. We chose to use the range representation of the scene $f(\theta,r)$ as the benchmark, assuming that the target and its associated $\rho(x,y)$ are unknown to the observer. It can be seen that the adaptive algorithm approaches low MSE faster than random sampling. Although the advantage tends to saturate when the number of measurements approaches the allowed maximum, the narrow error bars imply that the adaptive algorithm tends to be more stable when approaching convergence as the adaptive algorithm exploits the asymmetry of the target to select which measurements will most increase the information content. A histogram of the number of measurements required to reach 15% MSE further demonstrates the advantage of the adaptive algorithm over random sampling, as shown in Fig. 4(d). Although it happens that random sampling may converge faster than the adaptive algorithm in some cases, the adaptive algorithm is overall much more reliable and converges more consistently. This is essential when the target has unknown complexity and the number of measurements to be made must be predetermined.

It can also be noticed in Fig. 4(d), there was one outlier for the adaptive sampling that used all 60 measurements to reach the 15% MSE mark. The outlier is a consequence of the sampling constraints described earlier. Since ISAR has no multiplexing ability, each measurement is constrained to a single angle, and the algorithm must maximize the sum of $\mathbf{r}_{K+1}{}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$. Since the constraint may not allow the algorithm to choose the highest eigenvalue of $\boldsymbol{\Sigma}$ and associated eigenvector $\mathbf{r}_{K+1}$, the algorithm's more limited choices may be among lesser $\mathbf{r}_{K+1}$ values that may sometimes produce the same sum of $\mathbf{r}_{K+1}{}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$, and their corresponding angles may differ significantly. Consequently, the algorithm may sometimes choose an angle that does not aid convergence.

Choosing the next measurement angle to maximize variance necessarily minimizes the new entropy, as shown in Eq. (19). In other words, the algorithm chooses to maximize "new" information with each new measurement. To understand the selection process better, we show in Fig. 5(a)-5(d) four different stages of one simulated adaptive measurement, i.e. the 4th, 8th, 12th, and 16th adaptive measurement, respectively. The colored lines indicate the four measurement angles added in each sequence. Note that the early measurements are relatively evenly spaced in angle but that the later measurements have identified the thinner "eclipse" section for further scrutiny. These intuitive choices illustrate the algorithm's ability to recognize the part of the target with the most angle-dependent variation in signal.

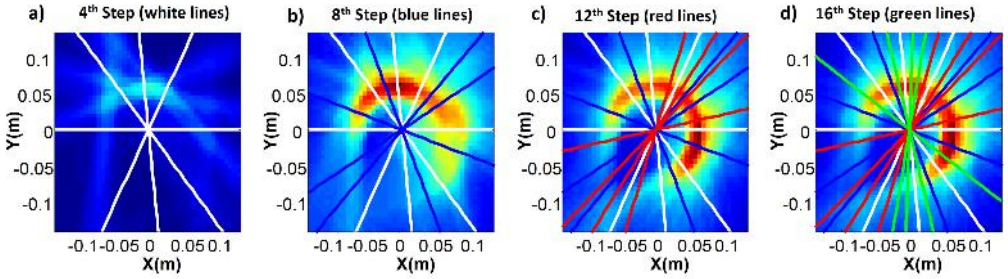

Fig. 5. (a)-(d). Results from the adaptive algorithm after 4, 8, 12, 16 measurements respectively. Lines with different colors indicate the measurements made during each stage, i.e., white lines are the first 4 measurements; blue lines are the 5th-8th measurements. These figures demonstrate the process of adaptive selection.

Are there conditions where adaptive sampling does not have a significant advantage over random sampling? Consider a cylindrical target for which all measurement angles return the same signal. In this case, each measurement would introduce the same amount of uncertainty,

Optics EXPRESS

and the adaptive algorithm would simply scan every angle. Thus, in scenes with cylindrical symmetry, the adaptive algorithm will not outperform the random sampling strategy.

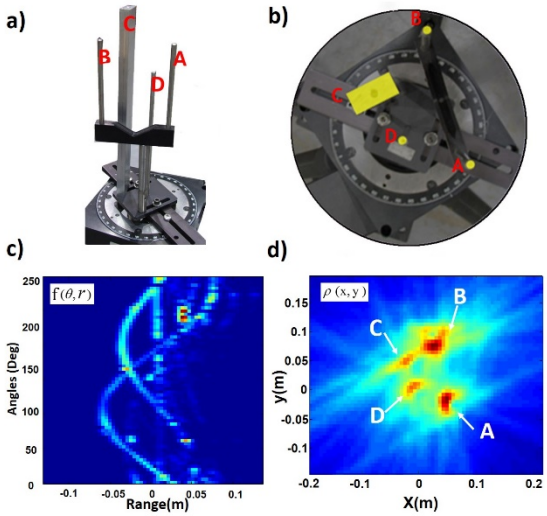

Fig. 6. Image of the target metal posts a) from the side and b) from above. c) The range signal measured at 300 angles separated at 0.1 degrees per measurement. d) The ISAR reconstruction via back propagation.

This can be observed in a second experiment involving a simpler target with cylindrical symmetry, composed of four metal posts that appear as point scatterers in the image. The configuration is shown in Figs. 6(a) and 6(b), and the range-angle data $f(\theta,r)$ is shown in Fig. 6(c) from which $\rho(x,y)$ is reconstructed in Fig. 6(d). Although the geometry of this target is simpler than the previous experiment, Fig. 7(a) reveals that the wavelet coefficients of the measurements in this second experiment are less sparse than that of the first experiment, in part because of the cylindrical symmetry. Figure 7(b) shows the reconstruction of the target with 35 measurements from the adaptive algorithm, which has an SSIM index of 0.917 compared to the benchmark in Fig. 6(d).

As predicted, the symmetry and non-sparsity impose difficulties for the adaptive algorithm. To compare the adaptive algorithm with the random sampling method, we again simulated 60 different experiments, each starting at a different measurement angle. As expected, the MSE progression of both methods shows that the adaptive algorithm did not outperform the random sampling algorithm as in Fig. 7(c). The two algorithms tend to approach a given MSE at the same rate. We believe this is due to the geometry of the target, for which most angles introduce a similar amount of uncertainty, so random selection is not an inferior methodology in this scenario. However, the adaptive algorithm still has the advantage of being more consistent, which is shown in Fig. 7(c) as the uncertainties associated with adaptive sampling are smaller. In addition, the histogram in Fig. 7(d) indicates how many measurements each algorithm took to achieve an MSE of 15%. The adaptive algorithm tends to achieve this goal in 45-50 measurements, whereas random sampling was less reliable by taking 35-55 measurements.

To characterize the advantage BCS provides, it is not enough to quantify how many fewer measurements were required to achieve a certain MSE. We must also consider the associated time and complexity introduced by the algorithm. According to the proposed method for selecting the next measurement, $\mathbf{r}_{K+1}{}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$ must be calculated for all remaining possible angles so that the best $\mathbf{r}_{K+1}$ may be selected, but the extra time for the matrix multiplication to take place is not always negligible. Two factors affect the calculation time: sampling in the

Optics EXPRESS

range domain, and sampling in the angular domain. When the range domain is heavily sampled with a large spectral bandwidth, $\mathbf{r}_{K+1}$ can be very large and the extra multiplication become a computational burden. Similarly, if the sampling in angular domain is very fine, there are too many possible $\mathbf{r}_{K+1}$ candidates. For example 360° sampled at 0.1° intervals produces 3600 choices for $\mathbf{r}_{K+1}$ at the beginning, dramatically increasing the time required for calculating all $\mathbf{r}_{K+1}{}^{\mathrm{T}}\boldsymbol{\Sigma}\mathbf{r}_{K+1}$ values. However, note that even when these computations slow the time for convergence, the stability and confidence of convergence remains a significant advantage for applying the adaptive BCS algorithm to ISAR imaging. It can be seen from the work presented here that regardless of scene complexity, the adaptive is more reliable than random sampling toward reaching a desired MSE.

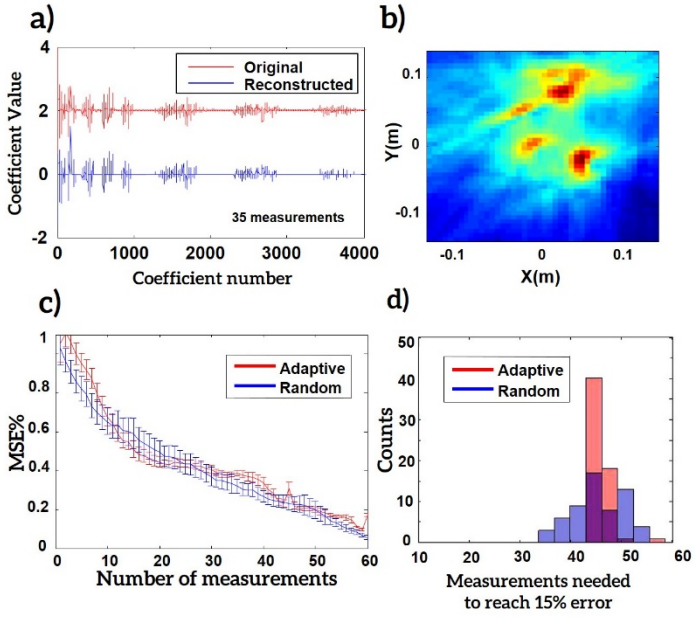

Fig. 7. a) The reconstruction of the wavelet coefficient of the signal after 35 measurements. b) The ISAR reconstruction via filtered back projection on the data with 35 measurements. c) The average MSE comparison of adaptive spatial sampling comparing to random spatial sampling for 60 simulated experiments. Error bars indicate the standard deviation of the MSE. d) Histogram of how many measurements the adaptive and random sampling has taken to reach 15% MSE.

4. Conclusion

We have presented an adaptive spatial sampling method for MMW inverse synthetic aperture radar (ISAR) imaging systems. The goal of adaptive imaging in ISAR for single pixel detector systems is to reduce the amount of mechanical scanning and the amount of data acquired in order to reduce acquisition time and minimize noise. We have implemented this adaptive algorithm based on the Bayesian compressive sensing algorithm developed by Carin et. al. [23], utilizing the full posterior estimation methodology to obtain optimized measurements. As compared to existing random sampling strategies, the advantage of this adaptive algorithm is the experimentally demonstrated ability to reconstruct scenes reliably without prior knowledge of the complexity of the scene.

Our results indicate that the complexity of the scene determines the degree to which this adaptive ISAR sampling strategy is superior to random sampling methods. Specifically, the advantage of this adaptive algorithm is diminished as the complexity or symmetry of the

Optics EXPRESS

scene increases and illumination from new directions provides little additional information. However, even in these scenarios the convergence is more consistently and reliably reached by adaptive methods.

Funding

U.S. Army Aviation and Missile Research, Development, and Engineering Center 100010185
Graduate School, Duke University 100006830.