



Millimeter wave imaging using sparse arrays

B. Mamandipoor^{a,*}, M. Fallahpour^b, A. Arbabian^b, U. Madhow^a

^aECE Department, University of California, Santa Barbara, CA, USA

^bEE Department, Stanford University, Stanford, CA, USA

ARTICLE INFO

Article history:

Received 4 February 2019

Revised 14 May 2019

Accepted 10 June 2019

Available online 11 June 2019

Keywords:

Mm-wave imaging

Sparse array

Grating lobe

Spatially-extended objects

Patch-based modeling

Sparse image reconstruction.

ABSTRACT

In this paper, we develop a theoretical framework for short-range millimeter (mm) wave radar imaging using a sparse array of monostatic elements, and validate it via experiments. The framework is a significant departure from classical radar, which largely focuses on long-range settings in which targets are well modeled as point scatterers. For sparse arrays, the point scatterer target model leads to grating lobes, and our central contribution is to demonstrate that a patch-based target model, suitably optimized for the sensor and scene geometry, suppresses such grating lobes. Key results include the following: (a) Characterizing the number of degrees of freedom (*DoF*) as a function of geometry, and showing that spatial undersampling (number of elements smaller than *DoF*) leads to grating lobes with the point target model; (b) showing that spatial aggregation via a patch-based dictionary suppresses grating lobes, and that patch size can be optimized based on estimation-theoretic criteria; (c) providing examples of the application, and adaptation, of patch-based dictionaries for sparse reconstruction.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

In this paper, we consider the problem of short-range radar imaging, motivated by the emergence of compact, low-cost radar sensors in the millimeter (mm) wave band [1,2], together with applications such as vehicular situational awareness [3] and gesture recognition [4]. Given a constraint on the form factor, we wish to design an array with a minimal number of sensors, and a minimal level of coordination among the sensors. We therefore consider a sparse array of monostatic sensors which are not synchronized in phase or frequency, and are loosely synchronized in time to the extent needed to avoid inter-sensor interference.

Classical radar theory and algorithms have been developed for long-range applications, in which targets are well modeled by point scatterers. For the sparse arrays considered here, the point target model leads to grating lobes. The central thesis of this paper is that a patch target model, suitably optimized for the sensor and scene geometry, suppresses such grating lobes. We establish this thesis by geometric and estimation-theoretic computations, and illustrate its application for sparse reconstruction techniques applied to experimental data acquired using a testbed in which a quasi-monostatic transceiver at 60 GHz (wavelength of $\lambda = 0.5$ cm) is

used to emulate (for static scenes) a two-dimensional (2D) array using a movable platform. These contributions are organized as follows.

1.1. Contributions

- (1) Using a Fresnel approximation, we show in Section 2 that the number of spatial degrees of freedom (*DoF*) available for imaging are limited by the spatial extent of the array and the scene. The result is analogous to the number of time-frequency degrees of freedom in a bandlimited channel [5,6], or the number of spatial degrees of freedom for Line-of-Sight MIMO communication [7]. Increasing the number of sensors beyond *DoF* can improve signal-to-noise ratio (*SNR*), but does not improve normalized measures of target discrimination.
- (2) In Section 3, we introduce the patch-based target model. We consider sparse arrays with number of elements smaller than *DoF*, for which conventional processing using a point target model is well known to lead to grating lobes; that is, point scatterers in spatially separated locations have highly correlated array responses. We note that at short ranges, targets are better modeled as a continuum of points, and introduce a dictionary in which each atom corresponds to a “patch,” or a continuum of point scatterers whose reflectivity is approximated as constant. We show that both grating lobes and side lobes are attenuated by such “spatial aggregation.”

* Corresponding author.

E-mail addresses: bmamandi@ece.ucsb.edu (B. Mamandipoor), mfallahpour@ieee.org (M. Fallahpour), arbabian@stanford.edu (A. Arbabian), madhow@ece.ucsb.edu (U. Madhow).

- (3) Experimental results demonstrating the effectiveness of matched filtering with respect to the patch basis functions for suppressing grating lobes are presented in Section 4, and are contrasted with standard imaging by matched filtering against point target basis functions.
- (4) In Section 5, we describe an estimation theoretic approach for tuning the patch size. The size of the grating lobes relative to the main lobe is related to the SNR threshold at which the Ziv-Zakai bound (ZZB) converges to the Cramér-Rao bound (CRB), and we choose the patch size based on when the SNR threshold for the sparse array matches that for an array with DoF elements.
- (5) In Section 6, we show that the patch model, in addition to helping with standard “matched filter” style imaging, also provides an effective basis for *sparse representation* of simple scenes. Specifically, if the scene reflectivity is low-pass, then patch-based basis functions are able to provide a parsimonious representation of the information in the scene.

1.2. Related work

The focus of this paper is to extract information from a *static scene* by reconstructing the spatial configuration of the objects and creating an image of the scene. This is quite different from the scenario in which the desired information resides in the temporal changes of the received signal, e.g., for recognizing *motion gestures* of a human hand [4,8]. Our patch model and sparse reconstruction approach falls within the framework of synthesis-based sparse signal representation, in which dictionary construction and image formation explicitly incorporate prior information about the sensing mechanism and the scene [9–11]. The patch primitive, for example, assumes that scene reflectivity is *low-pass* (i.e., it varies slowly in space), and the choice of patch size depends on the array and scene geometry, and the number of sensors.

Our approach for finding the maximally sparse representation of the scene in the dictionary of spatially extended objects falls within a general “sparse reconstruction” framework [12–14], which makes it possible to employ general-purpose ℓ_1 -regularized convex optimization methods [15]. However, we find that an extension of the Newtonized Orthogonal Matching Pursuit (NOMP) algorithm developed in [16,17], adapted here for sparse estimation in the dictionary of spatially extended objects, is more effective. Since our main goal here is to illustrate the efficacy of the patch model in suppressing grating lobes, detailed comparison between various sparse reconstruction algorithms is beyond our present scope; see [16] for extensive comparisons of performance and computational complexity in a more generic setting.

It is interesting to compare our system model to that employed in compressive sensing. Compressive sensing employs pseudorandom projections from a high-dimensional ambient space to a lower-dimensional observation space. In our setting, we have a low-dimensional measurement space by the physical use of fewer sensors, and consider a uniform rather than a random configuration. In each case, however, sparsity can be exploited for signal reconstruction or parameter estimation using a number of measurements which is much smaller than the ambient dimension, and similar algorithms can be employed, including the NOMP algorithm [16] that we adapt here. Thus, continuing advances in sparse reconstruction algorithms motivated by compressive sensing (e.g., [18–20]) could potentially be leveraged for further improving image reconstruction in our setting. This is beyond our present scope, but, as discussed in the conclusions, is an interesting topic for future work.

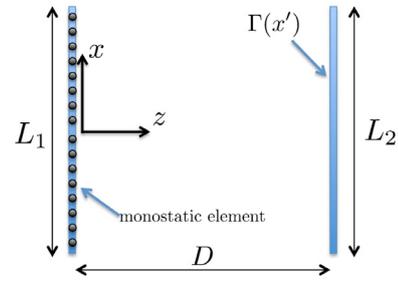


Fig. 1. Geometry of One-dimensional monostatic imaging.

This work builds on our prior conference papers [21–23], where we introduced some of the underlying concepts and provided some experimental results. We provide here a more complete theoretical treatment of designing patch-based dictionaries, and of using estimation-theoretic bounds for tuning the parameters of these models. We also present experimental results to demonstrate the applicability of the proposed object models for wideband three-dimensional imaging using a sparse array of antennas. Finally, we illustrate how the proposed patch-based dictionary can be combined with sparsity-aware image reconstruction schemes based on convex optimization and greedy pursuit, which aim to approximate the scene response by a linear combination of a few atoms from the dictionary of spatially extended objects.

2. Imaging geometry for the point scatterer model

We first identify the number of degrees of freedom (DoF) as a function of array and scene geometry, and then discuss how grating lobes occur when the number of elements is smaller than DoF .

2.1. Degrees of freedom

Fig. 1 depicts a one-dimensional aperture of length L_1 imaging a one-dimensional scene of length L_2 , with D denoting the distance between the aperture and the scene. (The DoF in a 2D system is a product of the DoF along each axis.) Consider the limit where there is a continuum of elements in the array, with x denoting the location of an arbitrary element. Let $\Gamma(x')$ denote the complex reflection coefficient for a point scatterer at location x' . The response at the aperture is governed by the Helmholtz wave equation (simplified by dropping space attenuation factors) [24], with the response at an element at location x given by

$$r(x) = \int_{-L_2/2}^{L_2/2} \Gamma(x') \xi(x', x) dx', \quad (1)$$

where

$$\xi(x', x) = e^{-j2kR(x', x)}, \quad (2)$$

denotes the space-variant impulse response of the system, $k = 2\pi/\lambda$ is the wavenumber, and $R(x', x) = \sqrt{D^2 + (x' - x)^2}$ is the path length from the transceiver location x to the point scatterer at location x' . For $D \gg L_1, L_2$, we can use a Fresnel approximation [25] (i.e., a first-order Taylor approximation) for the path length between the element at x and the scatterer at x' , to obtain

$$R(x', x) \approx D + \frac{(x' - x)^2}{2D}. \quad (3)$$

We note that the preceding approximation can be refined [26] when D is comparable to L_1 and L_2 , as it is for the nominal values in our simulations and experiments: $L_1 = L_2 = 15$ cm and $D = 30$ cm in our simulations. However, the Fresnel approximation

suffices for our present purpose of deriving high-level sparse array design principles. Plugging (3) into (1), we obtain

$$r(x) \approx e^{-j2kD} \int_{-L_2/2}^{L_2/2} \Gamma(x') e^{-j\frac{k}{D}(x'-x)^2} dx', \quad (4)$$

for $-L_1/2 \leq x \leq L_1/2$. The integral kernel here has the same form as that investigated by Slepian et al. [5,6] for timelimited and bandlimited functions (slightly adjusted by Fresnel quadratic phase masks [27]). The eigenfunctions are known to be prolate spheroidal wave functions. The eigenvalues remain approximately equal until a critical transition point ($\approx 2\frac{L_1L_2}{\lambda D}$), after which they rapidly decay to zero. Thus, we may define *DoF* as this critical number of nonzero eigenvalues:

$$DoF \approx 2\frac{L_1L_2}{\lambda D} + 1. \quad (5)$$

In our context, *DoF* is the maximum number of linearly independent measurements of the scene $\Gamma(x')$, $-L_2/2 \leq x' \leq L_2/2$, that can be achieved by an array of monostatic elements spanning an aperture of length L_1 . Increasing the number of elements beyond *DoF* can therefore only improve the SNR, not gather new information about the scene.

Similar arguments building on the work of Slepian et al. have been applied to determine *DoF* in various contexts, including diffraction-limited optics [28] and line-of-sight MIMO [7]. It is interesting to note that, compared to the MIMO communication system in [7], we gain a factor of two in *DoF* for a similar geometry due to the round-trip phase in a radar system. We next discuss how *DoF* determines the minimum number of elements in a sparse monostatic array required to avoid grating lobes.

2.2. Grating lobes

While we characterized *DoF* by considering an array with a continuum of elements, let us now consider an array with N elements with a uniform inter-element spacing of $d = L_1/(N-1)$, denoting by $x_n = nd$, $n = 0, 1, \dots, N-1$, the location of the n^{th} element. We now discuss when grating lobes arise for the point scatterer model, where a grating lobe corresponds to two spatially separated point scatterers having very similar array responses. To this end, consider two point scatterers located at x'_1 and x'_2 , and denote by \mathbf{r}_1 and \mathbf{r}_2 the corresponding $N \times 1$ array responses, respectively. Using the Fresnel approximation in (3), we obtain the response at n^{th} element,

$$\mathbf{r}_i[n] \approx e^{-j2kd} \Gamma_i e^{-j\frac{k}{D}(x'_i - x_n)^2}, \quad (6)$$

where $n = 0, 1, \dots, N-1$, and $i = 1, 2$. We now compute the correlation between the two responses. We have

$$\begin{aligned} \mathbf{r}_1^H \mathbf{r}_2 &\approx \sum_{n=0}^{N-1} \Gamma_1^* \Gamma_2 e^{j\frac{k}{D}(x'_1 - x_n)^2} e^{-j\frac{k}{D}(x'_2 - x_n)^2} \\ &= \Gamma_1^* \Gamma_2 e^{j\frac{k}{D}(x'_1{}^2 - x'_2{}^2)} \sum_{n=0}^{N-1} e^{j\frac{2k}{D}(x'_2 - x'_1)x_n}. \end{aligned} \quad (7)$$

It is convenient to define the normalized distance between the two point scatterers as

$$\delta \triangleq \frac{2k}{D}(x'_2 - x'_1)d = 4\pi \frac{(x'_2 - x'_1)L_1}{\lambda D}. \quad (8)$$

Note that $|\delta| \leq \frac{2kdL_2}{D} = 4\pi \frac{L_2d}{\lambda D}$ for a scene of extent L_2 . We can now compute the magnitude of the inner product between the responses as follows:

$$|\mathbf{r}_1^H \mathbf{r}_2| = |\Gamma_1| |\Gamma_2| \left| \frac{\sin(N\delta/2)}{\sin(\delta/2)} \right| = |\Gamma_1| |\Gamma_2| |\text{Dir}(\delta)|, \quad (9)$$

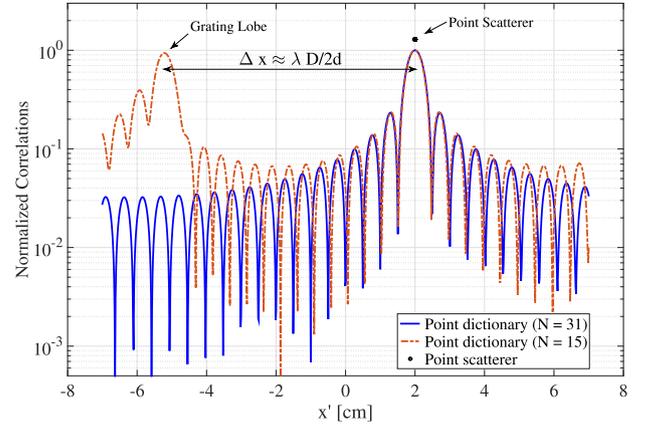


Fig. 2. Grating lobes for a sparse monostatic array. We do not have grating lobes for $N = 31 = DoF$, but grating lobes appear for a sparse array with $N = 15 < DoF$.

where $\text{Dir}(\delta) = \frac{\sin(N\delta/2)}{\sin(\delta/2)}$ is the well-known Dirichlet kernel. It is also easy to see that

$$\|\mathbf{r}_i\|^2 = |\Gamma_i|^2 N \quad i = 1, 2.$$

The normalized correlation between the two responses can now be computed to be

$$\rho = \frac{|\mathbf{r}_1^H \mathbf{r}_2|}{\|\mathbf{r}_1\| \|\mathbf{r}_2\|} = \frac{1}{N} |\text{Dir}(\delta)|. \quad (10)$$

The preceding normalized correlation $\rho = \rho(x'_1, x'_2)$ is also called the *ambiguity function*.

Number of array elements must be at least DoF to avoid grating lobes: The Dirichlet kernel is a periodic function with period 2π , hence the condition for avoiding aliasing, or grating lobes, corresponds to constraining the range of δ to be less than 2π , or $\frac{2kdL_2}{D} \leq 2\pi$. This yields $N \geq \frac{2L_1L_2}{\lambda D} + 1$ or $N \geq DoF$. That is, we avoid grating lobes if the number of array elements exceeds the degrees of freedom computed previously. For our nominal parameter values, we obtain that $N \geq 31$.

Predicting grating lobes with spatial undersampling: Spatial undersampling corresponds to $N < DoF$, and grating lobes can be easily predicted from the ambiguity function. Specifically, if we set $N = 15 < DoF = 31$, then we do expect to see a grating lobe when the normalized distance δ equals a nonzero integer multiple of 2π . Setting $\delta = \pm 2\pi$ in (8), we obtain that

$$\Delta x = |x'_2 - x'_1| = \frac{\lambda D}{2d} \approx 7.5 \text{ cm for } N = 15. \quad (11)$$

We illustrate the dependence of grating lobes on N in Fig. 2, which shows the magnitude of the normalized correlation for $x'_1 = 2$ cm fixed, and x'_2 varying in the interval $(-7.5, 7.5)$ cm, for $N = 15$ and $N = 31$. For $N = 15$, we have a grating lobe at $x_{gl} \approx -5.5$ cm, which is consistent with (11).

Increasing number of array elements beyond DoF does not improve ambiguity function: We now examine the behavior of the ambiguity function as we increase the number of array elements beyond *DoF*. The normalized distance is given by

$$\delta = 4\pi \frac{(x'_2 - x'_1)L_1}{(N-1)\lambda D},$$

so that

$$\sin(N\delta/2) \rightarrow \sin\left(2\pi \frac{(x'_2 - x'_1)L_1}{\lambda D}\right), \quad N \rightarrow \infty$$

and

$$N \sin \delta/2 \rightarrow 2\pi \frac{(x'_2 - x'_1)L_1}{\lambda D}, \quad N \rightarrow \infty$$

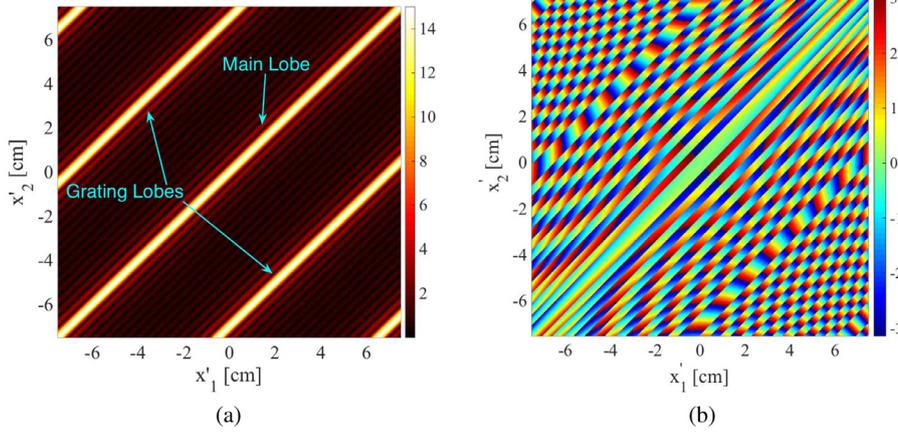


Fig. 3. (a) Magnitude of Dirichlet kernel ($N = 15$), (b) Phase of the product of SA kernel and Dirichlet kernel $H(x'_1, x'_2)\text{Dir}(\delta)$.

which implies that

$$\begin{aligned} \rho(x'_1, x'_2) &= \left| \frac{\sin(N\delta/2)}{N \sin(\delta/2)} \right| \rightarrow \left| \text{sinc}\left(2 \frac{(x'_2 - x'_1)L_1}{\lambda D}\right) \right| \\ &= \left| \text{sinc}\left(\text{DoF} \frac{(x'_2 - x'_1)}{L_2}\right) \right|, \quad N \rightarrow \infty \end{aligned} \quad (12)$$

where $\text{sinc}(x) \triangleq \frac{\sin(\pi x)}{\pi x}$. We note that the argument of the sinc function depends on DoF, but not on N . Thus, increasing the number of array elements beyond DoF does not improve the ambiguity function for locating a point scatterer in the scene. It only leads to an increase in effective SNR: the response energies $\|\mathbf{r}_i\|^2$ scale by N .

3. The patch model

We now show that grating lobes can be significantly attenuated by using a new patch primitive for scene reconstruction: a patch is a continuous collection of point scatterers adjacent to each other, with roughly constant reflection coefficient. This is well matched to scenes in which the reflectivity function is *spatially lowpass*.

Let us start with two arbitrary collections of point scatterers, Ψ_1 and Ψ_2 . Denote the corresponding array responses by $\tilde{\mathbf{r}}_1$ and $\tilde{\mathbf{r}}_2$, respectively. The correlation between the two responses can be computed as follows:

$$\begin{aligned} \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_2 &= \sum_{n=0}^{N-1} \int_{x'_1 \in \Psi_1} \Gamma^*(x'_1) \xi^*(x'_1, x_n) dx'_1 \int_{x'_2 \in \Psi_2} \Gamma(x'_2) \xi(x'_2, x_n) dx'_2 \\ &\approx \iint_{\substack{x'_1 \in \Psi_1 \\ x'_2 \in \Psi_2}} \Gamma^*(x'_1) \Gamma(x'_2) e^{j\frac{k}{2}(x'^2_1 - x'^2_2)} \sum_{n=0}^{N-1} e^{j\frac{2k}{2}(x'_2 - x'_1)x_n} dx'_2 dx'_1, \end{aligned} \quad (13)$$

using the Fresnel approximation as before. For normalized distance $\delta = \delta(x'_1, x'_2)$ defined as in (8), we see that this inner product can be written as

$$\tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_2 = c_1 \iint_{\substack{x'_1 \in \Psi_1 \\ x'_2 \in \Psi_2}} \Gamma^*(x'_1) \Gamma(x'_2) H(x'_1, x'_2) \text{Dir}(\delta) dx'_2 dx'_1, \quad (14)$$

where c_1 is a unit magnitude constant. We term $H(x'_1, x'_2) \triangleq e^{j\frac{k}{2}(x'^2_1 - x'^2_2)}$ the Spatial Aggregation (SA) kernel. The magnitude $|H(x'_1, x'_2)|$ is constant, while the phase is a nonlinear function of x'_1 and x'_2 . While the Dirichlet kernel depends on N , the SA kernel does not. Fig. 3 shows the magnitude of the Dirichlet kernel

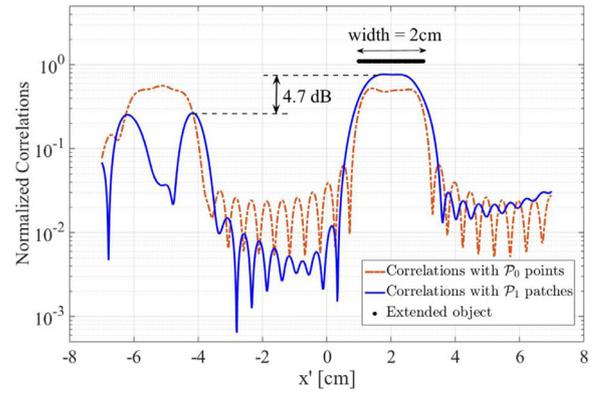


Fig. 4. Reduction of grating/side lobes by spatial aggregation.

as well as the phase of the product $H(x'_1, x'_2)\text{Dir}(\delta)$ for $x'_1, x'_2 \in (-7.5, 7.5)$ cm and $N = 15$. The key observation is that the phase of SA kernel is nearly constant across the main lobe, while exhibiting rapid variations across the grating lobes. The resulting incoherent integration in (14) results in significant suppression of grating lobes.

We now illustrate the impact of spatial aggregation for a dictionary of patches of width w cm for a 1D scene of length L_2 , specified as $\mathcal{P}_w \triangleq \{\Psi = [l - w/2, l + w/2] : l \in (-L_2/2, L_2/2)\}$. The dictionary of point scatterers is obtained by letting $w \rightarrow 0$. Fig. 4 shows, for $N = 15$, the magnitude of normalized correlations for a fixed patch of width 2 cm, setting $\Psi_1 = [1, 3]$ cm, against the dictionary of 1 cm patches, choosing $\Psi_2 \in \mathcal{P}_1$. We also plot the normalized correlation of the response to Ψ_1 against the dictionary of point scatterers \mathcal{P}_0 for comparison. We see that SA produces three main effects: (1) suppressing the grating lobe, (2) reducing the side lobe level, and (3) widening the main lobe. It is worth noting that widening of the main lobe is a natural consequence of a model aimed at representing an extended object with roughly constant reflectivity. Note that the correlation between atoms in a patch-based dictionary depends on N only through the Dirichlet kernel, and therefore inherits the latter's insensitivity to N when the number of elements exceeds DoF.

4. Experimental results

Our hardware testbed is a 60 GHz continuous-wave quasi-monostatic (transmit and receive antennas are slightly separated, but approximately appear to be co-located as viewed from the

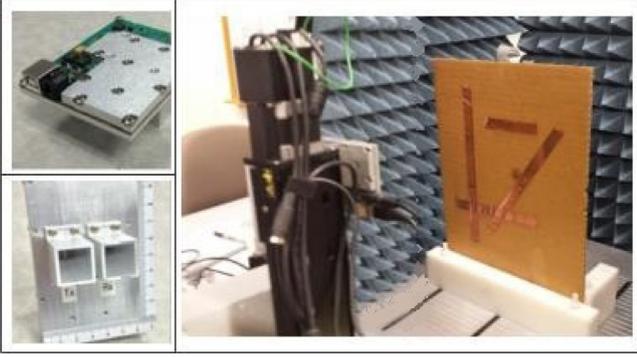


Fig. 5. Experimental data collection using 60 GHz quasi-monostatic radar system.

target) radar transceiver, equipped with dual high-gain horn antennas [29]. We use a mechanical platform to move the imager on a plane parallel to the scene, thereby emulating a 2D array of sensors. The travel distance of the imager is of the order of the form factor of a portable handheld device (covering a $15 \text{ cm} \times 15 \text{ cm}$ area).

4.1. 2D Reconstruction in the spatial domain

In this subsection, we consider two uniform planar array configurations: (1) Dense array of 30×30 elements (i.e. $d \approx \lambda = 0.5 \text{ cm}$), (2) sparse array of 15×15 elements (i.e. $d \approx 2\lambda = 1 \text{ cm}$). At each step of movement, the scene response is measured using a single frequency continuous-wave waveform at 60 GHz, and stored in a vector. We consider a scene of copper strips, shown in Fig. 5, that is placed parallel to the array at a distance $D = 30 \text{ cm}$. Fig. 6 shows the emulated 2D array configurations, covering a $15 \text{ cm} \times 15 \text{ cm}$ aperture.

The first reconstruction method that we consider is standard Synthetic Aperture Radar (SAR) processing, which may be viewed as matched filtering (MF) with respect to a point scatterer based signal model [30]. The second approach is based on computing the correlation of the measured response with the responses of $1.5 \text{ cm} \times 1.5 \text{ cm}$ square patches; that is, the collection of patches obtained by sliding a $1.5 \text{ cm} \times 1.5 \text{ cm}$ window over the entire scene. We refer to this approach as patch-based MF. Fig. 7 shows the results of point-based MF. We see that grating lobes lead to significant deterioration in image quality for the sparse array. Fig. 8 shows the result of patch-based MF. We see significant improvement in the image quality, in terms of suppressing the grating lobes and increasing the dynamic range, e.g. the horizontal

strip at the bottom of the scene becomes visible for the dense array deployment.

As we discussed in Section 2.2, the first grating lobe appears at distance $\Delta x \approx \frac{\lambda D}{2d}$ from the true point location. This corresponds to $\Delta x \approx 15 \text{ cm}$ and 7.5 cm , for our nominal geometry with $N = 30$ and $N = 15$ element arrays, respectively. The dependency of Δx on the wavelength λ , suggests that by incorporating a wideband signaling scheme, one can potentially suppress the grating lobe effects, due to the incoherency in the location of the grating lobes across different frequencies [31]. Next, we investigate the potential of wideband signaling in suppressing the grating lobes using experimental data for 3D image reconstruction.

4.2. Wideband SFCW signaling and 3D reconstruction

In this subsection, we investigate wideband Stepped-Frequency Continuous-Wave (SFCW) signaling using a 2D array of quasi-monostatic elements, followed by 3D image reconstruction techniques in the spatial domain. We consider two uniform planar array configurations: (1) Dense array of 50×50 elements (i.e. $d \approx 0.6\lambda = 0.3 \text{ cm}$), (2) sparse array of 17×17 elements (i.e. $d \approx 1.8\lambda = 0.9 \text{ cm}$). At each step of movement, the scene response is measured in discrete frequency steps, covering 55 to 60 GHz band in $N_\omega = 100$ uniformly spaced intervals. The overall response of scene is saved in a 3D matrix $f(x, y; \omega) \in \mathbb{C}^{N \times N \times N_\omega}$. We consider a scene of a plastic glove (with four fingers) filled with salt water, that is placed parallel to the array at a distance $D \approx 28 \text{ cm}$, as shown in Fig. 9.

We apply Matched Filter reconstruction technique across both space (aperture) and frequency (bandwidth) for any hypothesized location in the scene [32]. Fig. 10 shows the output of MF algorithm for the dictionary of point-scatterers, for both dense and sparse array configurations. The scene information is preserved for the dense array and we are able to identify three fingers and the palm of the hand in the image. The fourth finger, however, is not visible in the reconstructed image. This is due to the beam pattern of the horn antennas (low power illumination of the edges of the scene), as well as the small radar-cross-section of the finger, which lead to significant power loss for the backscattered electromagnetic wave. For the sparse array configuration, as shown in Fig. 10(b), point-based MF does not preserve the scene information due to the grating lobes. This indicates that the frequency diversity induced by wideband SFCW signaling is not sufficient for eliminating the grating lobes in the sparse monostatic array configuration. Next, we apply patch-based MF, which entails computing the correlation of the overall received response across space and frequency, with the dictionary of $1.5 \text{ cm} \times 1.5 \text{ cm}$ square patches at

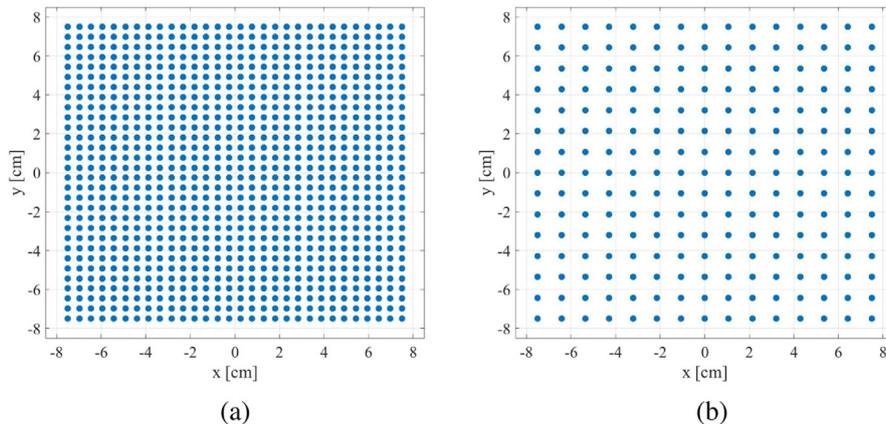


Fig. 6. Emulated 2D array configurations (a) dense 30×30 , and (b) sparse 15×15 elements, covering a $15 \text{ cm} \times 15 \text{ cm}$ aperture.

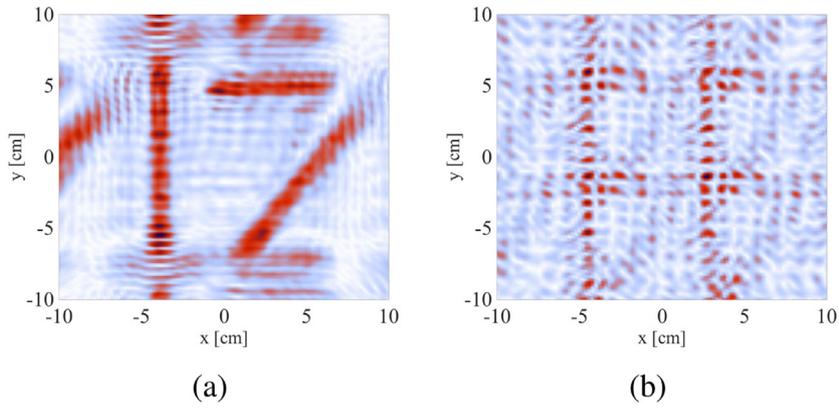


Fig. 7. Point-based MF (a) dense array, (b) sparse array.

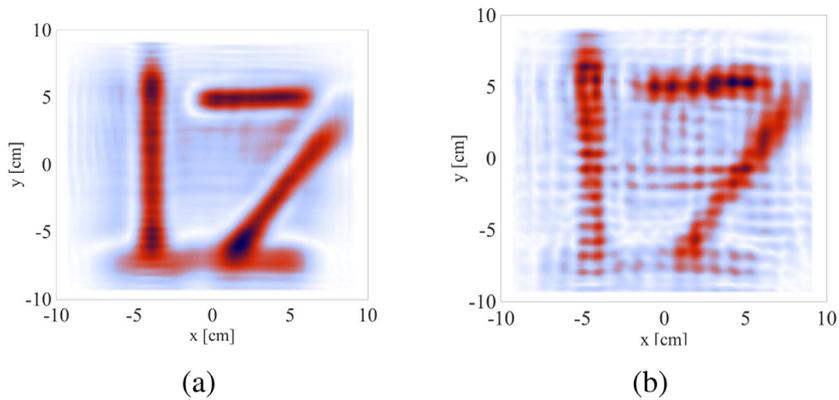


Fig. 8. Patch-based MF (a) dense array (b) sparse array.

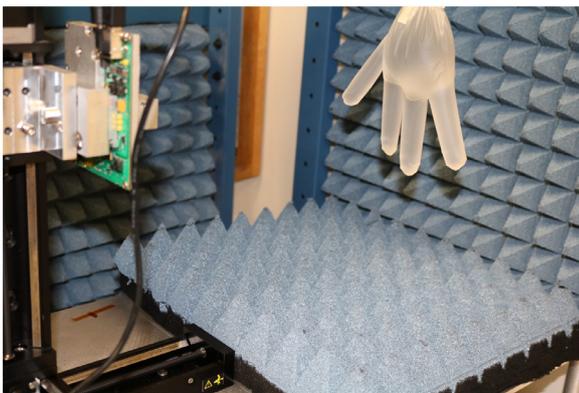


Fig. 9. Experimental data collection using SFCW radar system.

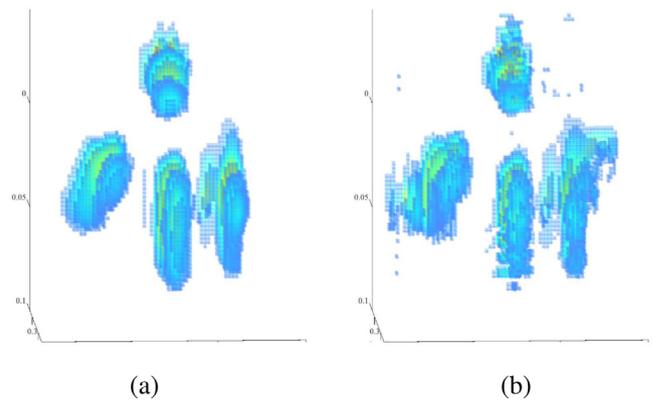


Fig. 11. Patch-based MF reconstruction of hand sample (a) dense array (b) sparse array.

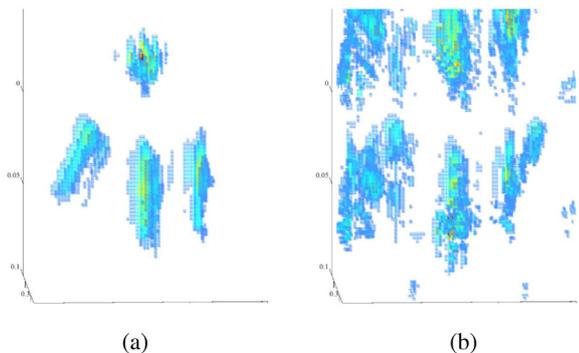


Fig. 10. Point-based MF reconstruction of hand sample (a) dense array (b) sparse array.

any hypothesized location in the scene. As shown in Fig. 11, the algorithm reduces grating lobes and preserves information in both dense and sparse array configurations.

5. Estimation-theoretic approach to choosing patch size

When we use a sparse array (number of elements smaller than DoF) with a patch-based dictionary, the choice of patch size trades off resolution (the main lobe is sharper with smaller patches) against grating lobe suppression (better with spatial aggregation over larger patches). Estimation-theoretic insight into this tradeoff can be obtained by considering the problem of locating a single atom of the dictionary in the presence of additive white Gaussian

noise (AWGN). We can divide the error for maximum likelihood (ML) estimation (or an approximation thereof) in such a problem into two-categories. The first is coarse-grained error, in which the estimated location is far away from the true location. In our context, this may be caused by a grating lobe or side lobe producing a larger peak than the main lobe in the presence of noise. The second is fine-grained error when the estimated location is close to the true location. This depends on the width of the main lobe. As long as the grating and side lobes are smaller than the main lobe, with high probability, we only incur fine-grained errors beyond some SNR threshold. However, for SNR smaller than this threshold, we may see substantially higher estimation error due to coarse-grained errors. This behavior of the ML estimator is well captured by the Ziv-Zakai bound (ZZB), which is a Bayesian bound (i.e., it assumes a prior distribution for the unknown parameter) on the mean squared error (MSE) of any estimator [33,34].

The ZZB captures the impact of both coarse- and fine-grained errors by relating the MSE in the estimation problem to the error probability in a hypothesis testing problem. As the SNR gets larger than a threshold, which we term the *ZZB threshold*, coarse-grained errors become less likely. After this point, the ML estimate, is close to the true parameter value [33] (i.e., it is in the “right bin”). Beyond the ZZB threshold, estimation performance is limited by main lobe ambiguity (i.e., with high probability, neither noise nor grating lobes can cause large estimation errors). In this regime, the ZZB is close to the Bayesian Cramér–Rao Bound (CRB), which operates under the assumption that we are in the right bin, and the MSE decays linearly with SNR on a log-log scale. We now sketch the derivation of the ZZB for our problem, and then discuss how we use the ZZB threshold as a design criterion for patch size.

5.1. Ziv-Zakai bound

Consider the problem of estimating the parameter $\theta \in \Theta \subset \mathbb{R}^K$ from noisy observations given by

$$\mathbf{y} = \mathbf{s}(\theta) + \mathbf{z}, \quad (15)$$

where $\mathbf{z} \sim \mathcal{CN}(0, \sigma^2 \mathbb{I}_N)$ is AWGN, and $\theta \sim P_\theta$. Let $\hat{\theta}(\mathbf{y})$ be an estimate of θ . The corresponding MSE matrix $\mathcal{M}(\hat{\theta})$ is defined by $\mathcal{M}_{m,n}(\hat{\theta}) \triangleq \mathbb{E}_{\mathbf{y},\theta}[(\hat{\theta}_m - \theta_m)(\hat{\theta}_n - \theta_n)]$. ZZB provides a lower bound on $\mathbf{a}^T \mathcal{M} \mathbf{a}$ for any $\mathbf{a} \in \mathbb{R}^K$, given by

$$\mathbf{a}^T \mathcal{M} \mathbf{a} \geq \frac{1}{2} \int_{h=0}^{\infty} \max_{\delta: \mathbf{a}^T \delta = h} \left\{ \int_{\phi \in \mathbb{R}^K} (P_\theta(\phi) + P_\theta(\phi + \delta)) P_{\min}(\phi, \phi + \delta) d\phi \right\} dh, \quad (16)$$

where $P_{\min}(\theta_1, \theta_2)$ is the minimum probability of error obtained from likelihood ratio test for the following hypothesis testing problem:

$$\begin{aligned} H_1 : \mathbf{y} &= \mathbf{s}(\theta_1) + \mathbf{z}, \Pr(H_1) = \frac{P_\theta(\theta_1)}{P_\theta(\theta_1) + P_\theta(\theta_2)}, \\ H_2 : \mathbf{y} &= \mathbf{s}(\theta_2) + \mathbf{z}, \Pr(H_2) = \frac{P_\theta(\theta_2)}{P_\theta(\theta_1) + P_\theta(\theta_2)}. \end{aligned} \quad (17)$$

Assuming uniform prior distribution P_θ , we have

$$P_{\min}(\theta_1, \theta_2) = Q\left(\frac{\|\mathbf{s}(\theta_1) - \mathbf{s}(\theta_2)\|_2}{\sqrt{2}\sigma}\right), \quad (18)$$

where $Q(\cdot)$ is the complementary cumulative distribution function (CCDF) of the standard Gaussian distribution $\mathcal{N}(0, 1)$. For our patch (or point) location estimation problem, the normalized measurement model boils down to

$$\mathbf{y} = e^{j\alpha} \tilde{\mathbf{r}}(l) + \mathbf{z}, \quad (19)$$

where α is an unknown phase parameter, and l denotes the location of the center of the patch object. The patch responses $\tilde{\mathbf{r}}(l) \triangleq \tilde{\mathbf{r}}(l)/\|\tilde{\mathbf{r}}(l)\|_2$ have been normalized to have unit energy values, hence $\text{SNR} = 1/\sigma^2$. Let α and l be independent random variables, uniformly distributed over the intervals $[0, 2\pi)$ and $(-L_2/2, L_2/2)$, respectively. The parameter vector $\theta \triangleq (\alpha, l)$. Setting $\mathbf{a} = [0, 1]^T$ gives us the ZZB bound on the MSE of the location estimation error denoted by MSE_l . It is easy to see that

$$\begin{aligned} & \max_{\delta: \mathbf{a}^T \delta = h} P_{\min}(\phi, \phi + \delta) \\ &= \max_{\delta: \mathbf{a}^T \delta = h} Q\left(\frac{\|\mathbf{s}(\phi) - \mathbf{s}(\phi + \delta)\|_2}{\sqrt{2}\sigma}\right) \\ &= \max_{\alpha \in [0, 2\pi)} Q\left(\frac{\|\tilde{\mathbf{r}}(l) - e^{j\alpha} \tilde{\mathbf{r}}(l+h)\|_2}{\sqrt{2}\sigma}\right) \\ &= Q\left(\sqrt{\frac{1 - |\tilde{\mathbf{r}}^H(l) \tilde{\mathbf{r}}(l+h)|}{\sigma^2}}\right). \end{aligned} \quad (20)$$

Therefore, the ZZB for estimating the location of patch objects of certain width reduces to the following,

$$\text{MSE}_l \geq \frac{1}{L_2} \times \int_{h=0}^{L_2} \int_{-L_2/2}^{L_2/2-h} Q\left(\sqrt{\frac{1 - |\tilde{\mathbf{r}}^H(l) \tilde{\mathbf{r}}(l+h)|}{\sigma^2}}\right) dl dh. \quad (21)$$

We evaluate (21) numerically for different array architectures and patch-based dictionaries \mathcal{P}_w .

5.2. Choosing patch size

Given a sparse array (i.e., number of elements smaller than *DoF*), our approach is to choose patch size so that the ZZB threshold does not exhibit significant deterioration relative to a system with as many array elements as the *DoF*.

Fig. 12 plots the ZZB for estimating the location of a point scatterer for different values of N , the number of array elements. We see that the bounds are indistinguishable for $N = 30$ and $N = 60$. This is a consequence of our previous observation that the normalized correlation function does not change much when we increase N beyond $\text{DoF} \approx 31$ elements (see Section 2). On the other hand, the grating lobes incurred when the number of elements is decreased to $N = 15$ are reflected in the corresponding ZZB curve: the ZZB threshold is about 13 dB larger than that for the dense arrays ($N = 30$ and 60). This significant increase in the ZZB threshold is because the multi-modal structure of the correlation function leads to a fundamental ambiguity in estimating the location of a single point scatterer.

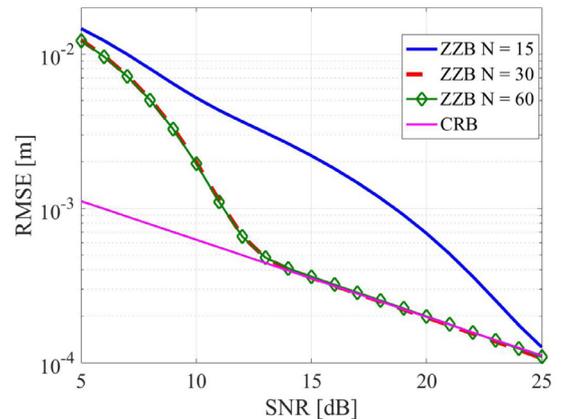


Fig. 12. ZZB for the dictionary of point-scatterers.

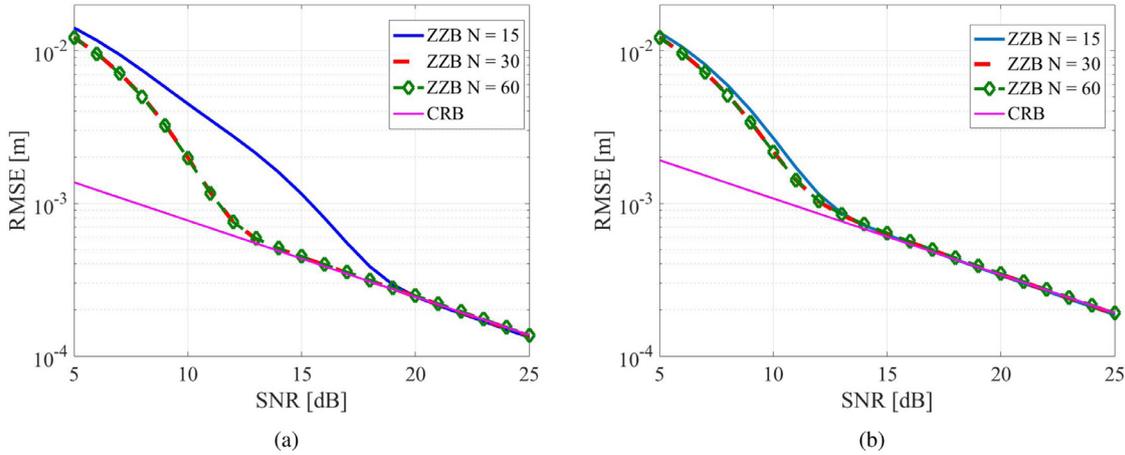


Fig. 13. ZB for estimating the location of a single patch in the dictionary of (a) patches of size $w = 0.5$ cm (b) patches of size $w = 1$ cm.

Now, consider the patch-based dictionary \mathcal{P}_w for patch width $w > 0$. For each candidate w , we compute the ZB for estimating the location of a single patch given that the size of the patch is known *a priori*. The gap between the ZB thresholds corresponding to the sparse array and a dense array is our metric for robustness against grating and side lobes. Fig. 13 shows that setting $w = 0.5$ cm leads to a 6 dB gap between the ZB thresholds for $N = 15$ and $N = 30$, whereas $w = 1$ cm eliminates this gap. Thus, for this particular scene geometry, array form factor, and number of elements ($N = 15$), a patch width of 1 cm is a better choice. This approach generalizes naturally to 2D arrays and more complicated parameterized dictionaries of spatially extended atoms.

In the next section, we investigate a new technique for image reconstruction that leverages the “sparse representation” of the scene in the dictionary of patches.

6. Sparsity-driven imaging based on patch dictionaries

Simple scenes usually admit a sparse representation in a dictionary of spatially-extended objects. For example, each finger in the image of the plastic glove in Section 4.2 can be approximated by a few concatenated patches. Such sparse representations have multiple potential advantages: (1) As pointed out in [9], sparse techniques for image formation can increase resolvability of targets, facilitate segmentation, and provide robustness to limitations in data quality and quantity; (2) sparse representations provide a framework for analyzing the scene at an “information rate” that is potentially significantly lower than the Nyquist rate [35], allowing minimalistic encoding of scene information; and (3) they may provide a basis for developing efficient algorithms for detecting scene changes via tracking the parameters of the estimated patches.

We formulate image formation as sparse reconstruction [12], assuming that we can approximate the scene response by a linear combination of a few atoms from a predefined dictionary. The response of each atom in the dictionary is represented by an N -dimensional vector $\mathbf{r} \in \mathbb{C}^N$. For a dictionary of ζ atoms, we construct a matrix $R \in \mathbb{C}^{N \times \zeta}$, whose columns are the responses of its atoms. That is, $R = [\mathbf{r}_1 \ \mathbf{r}_2 \ \dots \ \mathbf{r}_\zeta]$. The scene response is represented by

$$\mathbf{y} = R\mathbf{g} + \mathbf{z}, \quad (22)$$

where $\mathbf{g} \in \mathbb{C}^{\zeta \times 1}$ includes the complex gains corresponding to each atom, and $\mathbf{z} \sim \mathcal{CN}(0, \sigma^2 \mathbb{I}_N)$ is the AWGN. Let $\|\cdot\|_0$ denote the counting function (also known as ℓ_0 norm), which returns the number of nonzero elements of its input vector. Sparsity-driven imaging refers to the setting in which $\|\mathbf{g}\|_0 \ll \zeta$; that is, the scene admits a *sparse representation* in the constructed dictionary.

Our goal in sparse reconstruction framework is to find the *maximally sparse* representation of the scene, while allowing for some error tolerance $\epsilon \geq 0$ due to noise and modeling errors, by solving the following combinatorial optimization problem,

$$\underset{\mathbf{g}}{\text{minimize}} \ \|\mathbf{g}\|_0 \quad \text{subject to} \ \|\mathbf{R}\mathbf{g} - \mathbf{y}\|_2 < \epsilon. \quad (23)$$

Finding the exact solution of (23) without any constraints on the matrix R is known to be NP-hard [12]. Therefore, we resort to computationally tractable algorithms that generate approximate solutions. Specifically, in order to illustrate the utility of the patch model for sparse reconstruction, we focus on *Convex Relaxation* and *Greedy Pursuit*, two of the most popular techniques used in practice, which also offer certain theoretical guarantees of their performance [36,37]. We adapt the greedy pursuit technique presented here from [16], and refer the reader to the latter paper for a detailed comparison of performance-complexity tradeoffs between these two approaches.

6.1. Convex relaxation

A popular approach for sparse reconstruction is to replace the ℓ_0 norm in the optimization problem (23), with ℓ_1 norm, resulting in a convex optimization program,

$$\underset{\mathbf{g}}{\text{minimize}} \ \|\mathbf{g}\|_1 \quad \text{subject to} \ \|\mathbf{R}\mathbf{g} - \mathbf{y}\|_2 < \epsilon. \quad (24)$$

Note that the ℓ_1 norm is the closest convex function to the ℓ_0 function. It has been shown that under suitable conditions on the matrix R , and when the optimal \mathbf{g} is reasonably sparse, then this convex relaxation leads to the exact solution of the original problem in (23) [36]. One can also incorporate the ℓ_2 -error constraint in (24) as part of the objective function, yielding a scalarized dual-objective optimization program,

$$\underset{\mathbf{g}}{\text{minimize}} \ \frac{1}{2} \|\mathbf{R}\mathbf{g} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{g}\|_1, \quad (25)$$

where the regularization parameter $\lambda > 0$ balances the two objectives of minimizing the residual squared error (ℓ_2 term) and sparsity (ℓ_1 term). Increasing the value of λ typically leads to sparser solutions. Let $\lambda_{\max} \triangleq \|R^H \mathbf{y}\|_\infty$. Setting $\lambda > \lambda_{\max}$ leads to $\mathbf{g} = \mathbf{0}$ as the solution of (25). The formulation in (25) is also known as least absolute shrinkage and selection operator (LASSO), which was first introduced in the context of feature selection [15]. We apply LASSO to reconstruct the scene of copper strips (Fig. 5), given the sparse 15×15 element array configuration (described in Section 4.1). Fig. 14 shows LASSO outputs for different values of λ , when the matrix R is constructed based on the dictionary of $1.5 \text{ cm} \times 1.5 \text{ cm}$ square patches. The color of each patch

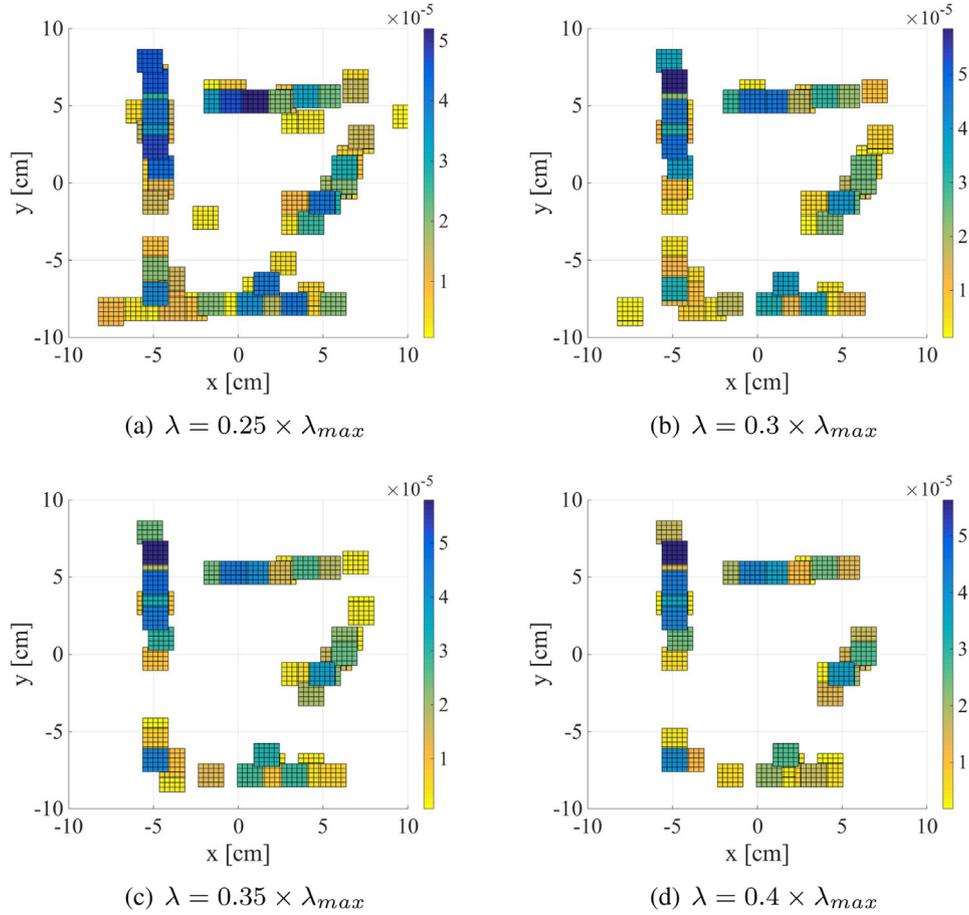


Fig. 14. Sparse reconstruction of the scene of copper strips (shown in Fig. 5), using LASSO for different values of λ .

is proportional to its gain. We see that increasing λ leads to sparser solutions, while preserving most of the information in the scene. However, due to the modeling errors in approximating the scene by square patches of a fixed size (with constant reflectivity across each patch), we see significant overlap of multiple patches trying to explain the reflectivity of the actual scene. As we show next, greedy approaches to sparse reconstruction work better.

6.2. Greedy pursuit

Another approach for finding the maximally sparse representation of the scene in the dictionary of patches is based on Greedy Pursuit [12]. We adapt here a recently developed iterative greedy algorithm named Newtonized Orthogonal Matching Pursuit (NOMP) [16,17], which is a generalization of the well-known Orthogonal Matching Pursuit (OMP) [37,38] to a continuously parametrized overcomplete basis using Newton refinements. Such greedy iterative approaches (e.g., OMP and NOMP) are particularly attractive due to their low computational complexity and ease of implementation [12].

In the original presentation of the NOMP algorithm [16,17], it is applied to estimation of frequencies in a mixture of sinusoids. However, we can apply the algorithm for sparse approximation in any continuously parametrized overcomplete dictionary. We now present a description of how NOMP can be employed for sparse recovery in our radar imaging problem. Let us first discuss the estimation of a single patch, and then build upon it to generalize to the estimation of a scene containing a mixture of patches.

6.2.1. Single patch

For a single patch, the measurement model (22) reduces to

$$\mathbf{y} = \mathbf{g}\mathbf{r}(l, w) + \mathbf{z}, \quad (26)$$

where $\mathbf{r}(l, w)$ denotes the response of a patch object of width w with its center located at l . Note that in a realistic scenario the patch object might not lie on the discrete grid identified by the columns of R , i.e., $\mathbf{r}(l, w) \notin \text{col}(R)$, where $\text{col}(R)$ is the column space of R . This is known as “off-grid effect” or basis mismatch [39] which degrades the performance of reconstruction schemes significantly. NOMP avoids the basis mismatch problem by searching over the continuum using Newton-based update steps. The maximum likelihood (ML) estimate of the gain g and patch parameters w, l are obtained by minimizing the residual power $\|\mathbf{y} - \mathbf{g}\mathbf{r}(l, w)\|_2^2$, or equivalently by maximizing the function

$$\mathcal{G}(g, l, w) = 2\Re\{\mathbf{y}^H \mathbf{g}\mathbf{r}(l, w)\} - |g|^2 \|\mathbf{r}(l, w)\|_2^2. \quad (27)$$

In this section, we assume a fixed w , and try to optimize (27) for g and l . The extension of the algorithm for optimizing w is straightforward, and is briefly discussed in Section 6.3. Directly optimizing (27) over all gains and locations is difficult. Therefore, NOMP adopts a two step procedure: (i) coarse detection over the discrete grid identified by the columns of R , (ii) iteratively refining the gain and location estimates.

For any given location l , the gain that maximizes (27) is given by $\hat{g} = (\mathbf{r}^H(l, w)\mathbf{y})/\|\mathbf{r}(l, w)\|_2^2$. Substituting \hat{g} in (27) yields the generalized likelihood ratio test (GLRT) cost function for estimating the patch location by

$$\bar{\mathcal{G}}(l, w) = |\mathbf{r}^H(l, w)\mathbf{y}|/\|\mathbf{r}(l, w)\|_2^2. \quad (28)$$

The GLRT estimate of l is derived by maximizing $\bar{G}(l, w)$. We use this observation to find a coarse estimate of the patch location over a predefined discrete grid, $\mathcal{L}_c \triangleq \{-L_2/2 + n(\frac{L_2}{n_{\text{points}} - 1}) : n = 0, 1, \dots, n_{\text{points}} - 1\}$, namely,

$$\hat{l}_c = \arg \max_{l \in \mathcal{L}_c} \bar{G}(l, w), \quad (29)$$

along with the corresponding gain given by $\hat{g}_c = (\mathbf{r}^H(\hat{l}_c, w)\mathbf{y}) / \|\mathbf{r}(\hat{l}_c, w)\|_2^2$. The refinement steps further enhances these estimates by searching over the continuum.

Refinement: The inputs to the refinement step are the current estimates of the patch parameters, denoted by $\{\hat{g}, \hat{l}\}$. The Newton step for location refinement is given by

$$\hat{l} = \hat{l} - \dot{G}(\hat{g}, \hat{l}, w) / \ddot{G}(\hat{g}, \hat{l}, w), \quad (30)$$

where

$$\dot{G}(\hat{g}, \hat{l}, w) = \Re\{(\mathbf{y} - \hat{g}\mathbf{r}(\hat{l}, w))^H \hat{g}(\mathbf{d}\mathbf{r}(\hat{l}, w)/d\mathbf{l})\}, \quad (31)$$

$$\ddot{G}(\hat{g}, \hat{l}, w) = \Re\{(\mathbf{y} - \hat{g}\mathbf{r}(\hat{l}, w))^H \hat{g}(d^2\mathbf{r}(\hat{l}, w)/d\mathbf{l}^2)\} - |\hat{g}|^2 \|\mathbf{d}\mathbf{r}(\hat{l}, w)/d\mathbf{l}\|_2^2. \quad (32)$$

Since we want to maximize the (generally non-concave) function $\bar{G}(g, l, w)$, we apply the Newton refinement step only when the function is locally concave, i.e., $\ddot{G}(\hat{g}, \hat{l}, w) < 0$. These refinement steps provide the flexibility of searching for a patch object in the “neighborhood” of our current estimated patch which can better explain the observed measurements. Note that after refinement the estimated patch may not lie on \mathcal{L}_c grid, thus avoiding the off-grid-effect of discretizing the parameter space.

6.2.2. Multiple patches

The response to a scene with \mathcal{K} patches is modeled as

$$\mathbf{y} = \sum_{i=1}^{\mathcal{K}} g_{\alpha_i} \mathbf{r}_{\alpha_i} + \mathbf{z}, \quad (33)$$

where $g_{\alpha_i} \in \mathbb{C}$ denotes the complex gain for i th patch. Let Λ be the set of detected patches. In each iteration of the algorithm, an atom that yields the greatest improvement in the approximation quality is identified and added to Λ . After that, a *cyclic refinement* step is applied to all of the atoms in Λ (i.e., to all of the atoms that have been estimated in the previous iteration), therefore giving them a chance to re-evaluate their estimates to incorporate the effect of the newly detected atom. We do not make *a priori* assumptions on the number of patches. Rather, the stopping criterion is based on the relative energy reduction of the residual signal (i.e., the portion of the signal not explained by the currently estimated set of patches). We stop looking for further patches when the relative energy reduction of the residual goes below a threshold, denoted by ϵ . The iterative sparse reconstruction algorithm is stated as follows:

- (1) Let $\mathbf{q}_0 = \mathbf{y}$, $\varepsilon_{q_0} = \|\mathbf{q}_0\|^2$, and loop counter $i \leftarrow 1$.
- (2) Find $\lambda = \arg \max_{\beta} \{ \frac{|\mathbf{q}_{i-1}^H \mathbf{r}_{\beta}|}{\|\mathbf{q}_{i-1}\| \|\mathbf{r}_{\beta}\|} : \mathbf{r}_{\beta} \in R \}$. Set $g_{\lambda} = \frac{(\mathbf{r}_{\lambda}^H \mathbf{q}_{i-1})}{\|\mathbf{r}_{\lambda}\|^2}$, and update $\Lambda \leftarrow \Lambda \cup \{(\lambda, g_{\lambda})\}$.
- (3) Cyclicly refine centers/gains for all patches in Λ .
- (4) Update all gains by least squares for best approximation of \mathbf{y} with the atoms chosen so far.
- (5) Find the new residual $\mathbf{q}_i = \mathbf{y} - \sum_{\lambda \in \Lambda} g_{\lambda} \mathbf{r}_{\lambda}$, and compute its energy $\varepsilon_{q_i} = \|\mathbf{q}_i\|^2$.
- (6) If $|\varepsilon_{q_i} - \varepsilon_{q_{i-1}}|/\varepsilon_{q_0} > \epsilon$, then set $i \leftarrow i + 1$, and go back to Step 2, otherwise, declare Λ as the output of the program.

Fig. 15 shows NOMP-based reconstruction of the same scene of copper strips (Fig. 5) in our earlier results, with a sparse

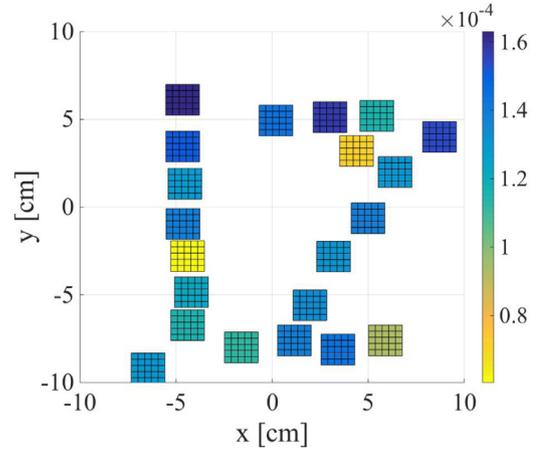


Fig. 15. Image reconstruction using NOMP algorithm in the sparse array configuration. The patch sizes are fixed (1.5cm \times 1.5cm), and we only refine the location of patches.

array with only 15 \times 15 equi-spaced emulated array elements. We employ a dictionary R with 1.5 cm \times 1.5 cm square patches. Despite the sparsity of the array, the NOMP-based reconstruction is able to detect the horizontal strip at the bottom of the scene, which is a significant improvement over the MF processing results in Figs. 7(b) and 8(b). Moreover, in comparison to the LASSO output in Fig. 14, we see that NOMP is more immune to the modeling errors, and is able to generate a sparse approximation of the scene with minimal overlap of the detected patches.

6.3. Dynamically adapted dictionaries

In our construction of patch-based dictionaries, we have explicitly incorporated prior information about the array geometry, as well as prior assumptions about scene characteristics (e.g., spatially lowpass reflectivity). Such adaptation of the dictionary to both sensing geometry and scene characteristics can be generalized to other settings (e.g., for non-uniform arrays). One possible approach for scaling this approach to more complex scenes without using an excessively large number of atoms is to start with sparse reconstruction with a base dictionary, and then to adapt the selected atoms on the fly. We report here on some experiments based on a natural extension of the NOMP algorithm for adapting dictionary parameters.

Consider the dictionary of fixed-size square patches employed in Section 6.2. We now augment it by allowing modification of patch sizes to better approximate the response to a given scene. This is easily accomplished by refining the width of patches along with their centers and gains in Step 3 of the NOMP algorithm. Fig. 16 shows the output of NOMP where we start with the dictionary of 1.5 cm \times 1.5 cm patches as the base, and then refine the sizes of the detected patches throughout the reconstruction process.

As another example, consider a dictionary of circular patches, using a collection of 1cm radius atoms as the base dictionary, but allowing for both center and radius refinements. An advantage of this circular dictionary is that the spatial size of the atoms is controlled by a single parameter (radius), hence dynamic adaptation is more computationally efficient than for a dictionary of square patches. Fig. 17 illustrates that sparse reconstruction with this dynamic dictionary does capture the structure of the scene, including the horizontal strip at the bottom of the scene.

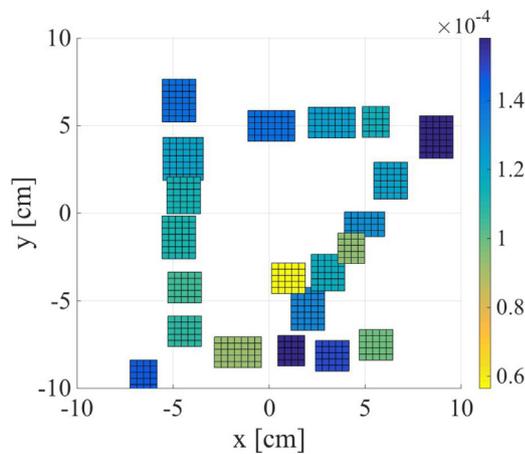


Fig. 16. Image reconstruction using NOMP algorithm in the sparse array configuration. We refine both the location and size of rectangular patches.

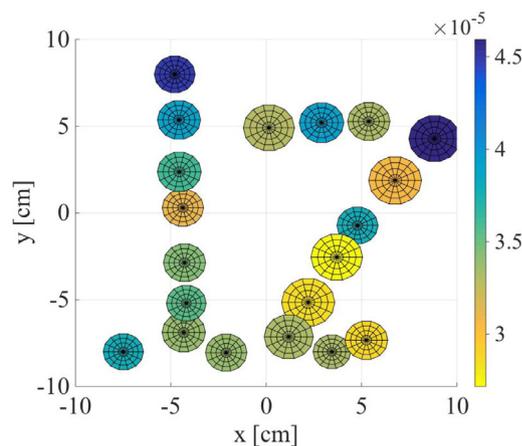


Fig. 17. Image reconstruction using NOMP algorithm in the sparse array configuration. We refine both the location and size of circular patches.

7. Conclusions

We have shown that short-range radar imaging requires new models that account for scene and transceiver geometry, as well as the number of transceiver elements. The number of degrees of freedom, DoF , in this setting depends on sensor and scene geometry, and using a number of elements smaller than DoF leads to grating lobes with the classical point scatterer model. The patch-based models introduced here suppress grating lobes, while being compatible with the spatially lowpass nature of typical scenes. The patch primitive can be employed for spatial matched filtering, as well as for sparse reconstruction.

While our preliminary results for sparse reconstruction are promising, enhancing the reconstruction leveraging continuing algorithmic advances (typically motivated by compressive sensing) is an interesting area for future research. In particular, it is of interest to explore how to bridge the gap between the noisier, but continuous, matched filter based reconstruction (Fig. 8(b)), and the gaps exhibited by the cleaner sparse reconstruction, even after dynamic adaptation (Figs. 16 and 17). Potential enhancement in performance by changing the sensor configuration (e.g., from a uniform to a random, or optimized, spacing) is also an interesting area of investigation. It is worth mentioning that for the special case where we construct the dictionary of patches by spatial translations of a single patch (that is fixed in shape and size), patch-based MF is equivalent to applying a low-pass filter (LPF) in the spatial-

frequency domain. This opens up another important area for future work for understanding the fundamental imaging modes in short-range settings in the spatial-frequency domain (preliminary results appear in [40]). Finally, from a practical standpoint, adding Doppler information to scene reconstruction with a patch model is an important area (see [41] for preliminary work with a point scatterer model).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA, and by the National Science Foundation under grants CNS-1518812 and CNS-1518632. The authors gratefully acknowledge Dr. Karam Noujeim of Anritsu and Greg Malysa for help with the mm-wave imaging and measurement setup.

References

- [1] S. Ahmed, A. Schiessl, F. Gumbmann, M. Tiebout, S. Methfessel, L. Schmidt, Advanced microwave imaging, *Microw. Mag. IEEE* 13 (6) (2012) 26–43, doi:10.1109/MMM.2012.2205772.
- [2] A. Arbabian, S. Callender, S. Kang, M. Rangwala, A. Niknejad, A 94 GHz mm-wave-to-baseband pulsed-radar transceiver with applications in imaging and gesture recognition, *Solid-State Circuits IEEE J.* 48 (4) (2013) 1055–1071, doi:10.1109/JSSC.2013.2239004.
- [3] M. Steinhauer, H.-O. Ruob, H. Irion, W. Menzel, Millimeter-wave-radar sensor based on a transceiver array for automotive applications, *IEEE Trans. Microw. Theory Tech.* 56 (2) (2008) 261–269.
- [4] J. Lien, N. Gillian, M.E. Karagozler, P. Amihoud, C. Schwesig, E. Olson, H. Raja, I. Poupyrev, Soli: ubiquitous gesture sensing with millimeter wave radar, *ACM Trans. Graphics (TOG)* 35 (4) (2016) 142.
- [5] D. Slepian, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty – i, *Bell Labs Tech. J.* 40 (1) (1961) 43–63.
- [6] H.J. Landau, H.O. Pollak, Prolate spheroidal wave functions, fourier analysis and uncertainty – iii: the dimension of the space of essentially time-and band-limited signals, *Bell Syst. Tech. J.* 41 (4) (1962) 1295–1336.
- [7] E. Torkildson, U. Madhoo, M. Rodwell, Indoor millimeter wave mimo: feasibility and performance, *IEEE Trans. Wirel. Commun.* 10 (12) (2011) 4150–4160.
- [8] C. Li, Z. Peng, T.-Y. Huang, T. Fan, F.-K. Wang, T.-S. Horng, J.-M. Muñoz-Ferreas, R. Gómez-García, L. Ran, J. Lin, A review on recent progress of portable short-range noncontact microwave radar systems, *IEEE Trans. Microwave Theory Tech.* (2017).
- [9] M. Cetin, I. Stojanovic, N. Onhon, K. Varshney, S. Samadi, W. Karl, A. Willsky, Sparsity-driven synthetic aperture radar imaging: reconstruction, autofocusing, moving targets, and compressed sensing, *Signal Process. Mag. IEEE* 31 (4) (2014) 27–40, doi:10.1109/MSP.2014.2312834.
- [10] S. Samadi, M. Cetin, M. Masnadi-Shirazi, Sparse representation-based synthetic aperture radar imaging, *Radar Sonar Navig. IET* 5 (2) (2011) 182–193, doi:10.1049/iet-rsn.2009.0235.
- [11] S. Samadi, M. Cetin, M.A. Masnadi-Shirazi, Multiple feature-enhanced SAR imaging using sparsity in combined dictionaries, *IEEE Geosci. Remote Sens. Lett.* 10 (4) (2013) 821–825, doi:10.1109/LGRS.2012.2225016.
- [12] J. Tropp, S. Wright, Computational methods for sparse solution of linear inverse problems, *Proc. IEEE* 98 (6) (2010) 948–958, doi:10.1109/JPROC.2010.2044010.
- [13] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, *SIAM Rev.* 43 (1) (2001) 129–159.
- [14] S. Foucart, M.-J. Lai, Sparsest solutions of underdetermined linear systems via ℓ_q -minimization for $0 < q \leq 1$, *Appl. Comput. Harmonic Anal.* 26 (3) (2009) 395–407.
- [15] R. Tibshirani, Regression shrinkage and selection via the lasso, *J. R. Stat. Soc. Series B* (1996) 267–288.
- [16] B. Mamandipoor, D. Ramasamy, U. Madhoo, Newtonized orthogonal matching pursuit: frequency estimation over the continuum, *IEEE Trans. Signal Process.* 64 (19) (2016) 5066–5081.
- [17] B. Mamandipoor, D. Ramasamy, U. Madhoo, Frequency estimation for a mixture of sinusoids: anear-optimal sequential approach, in: *3rd IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2015.
- [18] Y. Gu, N.A. Goodman, Information-theoretic compressive sensing kernel optimization and Bayesian cramer-Rao bound for time delay estimation, *IEEE Trans. Signal Process.* 65 (17) (2017) 4525–4537, doi:10.1109/TSP.2017.2706187.

- [19] S. Liu, Y.D. Zhang, T. Shan, R. Tao, Structure-aware Bayesian compressive sensing for frequency-hopping spectrum estimation with missing observations, *IEEE Trans. Signal Process.* 66 (8) (2018) 2153–2166, doi:10.1109/TSP.2018.2806351.
- [20] F. Li, S. Hong, Y. Gu, L. Wang, An optimization-oriented algorithm for sparse signal reconstruction, *IEEE Signal Process. Lett.* 26 (3) (2019) 515–519, doi:10.1109/LSP.2019.2897458.
- [21] B. Mamandipoor, G. Malysa, A. Arbabian, U. Madhow, K. Noujeim, 60 GHz synthetic aperture radar for short-range imaging: theory and experiments, in: *Signals, Systems and Computers, 2014 48th Asilomar Conference on*, IEEE, 2014, pp. 553–558.
- [22] B. Mamandipoor, M. Fallahpour, G. Malysa, K. Noujeim, U. Madhow, A. Arbabian, Spatial-domain technique to overcome grating lobes in sparse monostatic mm-wave imaging systems, in: *Microwave Symposium (IMS), 2016 IEEE MTT-S International*, IEEE, 2016, pp. 1–4.
- [23] B. Mamandipoor, A. Arbabian, U. Madhow, New models and super-resolution techniques for short-range radar: theory and experiments, in: *2016 Information Theory and Applications Workshop (ITA)*, 2016, pp. 1–7, doi:10.1109/ITA.2016.7888138.
- [24] D.K. Cheng, et al., *Field and Wave Electromagnetics*, Pearson Education India, 1989.
- [25] W. Southwell, Validity of the fresnel approximation in the near field, *JOSA* 71 (1) (1981) 7–14.
- [26] B. Mamandipoor, A. Arbabian, U. Madhow, Geometry-constrained degrees of freedom analysis for imaging systems: monostatic and multistatic, (2018) arXiv:1711.03585v2.
- [27] H. Lee, *Acoustical Sensing and Imaging*, CRC Press, 2016.
- [28] A. Thaning, P. Martinsson, M. Karelin, A.T. Friberg, Limits of diffractive optics by communication modes, *J. Optics A* 5 (3) (2003) 153.
- [29] K. Noujeim, G. Malysa, A. Babveyh, A. Arbabian, A compact nonlinear-transmission-line-based mm-wave sfcw imaging radar, in: *44th European Microwave Conference (EuMC)*, 2014, pp. 1766–1769, doi:10.1109/EuMC.2014.6986799.
- [30] R. Bamler, Principles of synthetic aperture radar, *Surv. Geophys.* 21 (2–3) (2000) 147–157.
- [31] M.P. Hayes, P.T. Gough, Broad-band synthetic aperture sonar, *IEEE J. Oceanic Eng.* 17 (1) (1992) 80–94.
- [32] S.S. Ahmed, A. Schiessl, L.-P. Schmidt, Illumination properties of multistatic planar arrays in near-field imaging applications, in: *Radar Conference (EuRAD), 2010 European*, IEEE, 2010, pp. 29–32.
- [33] H. Van Trees, K. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley, 2007.
- [34] K.L. Bell, Y. Ephraim, H.L.V. Trees, Explicit Ziv–Zakai lower bound for bearing estimation, *IEEE Trans. Signal Process.* 44 (11) (1996) 2810–2824.
- [35] R. Baraniuk, P. Steeghs, Compressive radar imaging, in: *2007 IEEE Radar Conference*, 2007, pp. 128–133, doi:10.1109/RADAR.2007.374203.
- [36] E.J. Candès, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inf. Theory* 52 (2) (2006) 489–509.
- [37] J. Tropp, A. Gilbert, Signal recovery from random measurements via orthogonal matching pursuit, *IEEE Trans. Inf. Theory* 53 (12) (2007) 4655–4666, doi:10.1109/TIT.2007.909108.
- [38] S.G. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, *IEEE Trans. Signal Process.* 41 (12) (1993) 3397–3415.
- [39] Y. Chi, L.L. Scharf, A. Pezeshki, A.R. Calderbank, Sensitivity to basis mismatch in compressed sensing, *IEEE Trans. Signal Process.* 59 (5) (2011) 2182–2195.
- [40] B. Mamandipoor, U. Madhow, A. Arbabian, 2d mm-wave imaging based on singular value decomposition, *Microwave Symposium (IMS), 2019 IEEE MTT-S International*, IEEE, In press, 2019.
- [41] A. Gupta, U. Madhow, A. Arbabian, Super-resolution in position and velocity estimation for short-range mm-wave radar, in: *2016 50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 1144–1148, doi:10.1109/ACSSC.2016.7869550.