60 GHz Synthetic Aperture Radar for Short-Range Imaging: Theory and Experiments

B. Mamandipoor(1), G. Malysa(2), A. Arbabian(2), U. Madhow(1) and K. Noujeim(3)

(1) ECE Department, University of California, Santa Barbara, CA, USA
(2) EE Department, Stanford University, Stanford, CA, USA
(3) Microwave Measurements Division, Anritsu Co., Morgan Hills, CA, USA

Emails: {bmamandi, madhow}@ece.ucsb.edu
{gmalysa, arbabian}@stanford.edu, Karam.Noujeim@anritsu.com

Abstract—We report on experimental results, and associated theory, for a 60 GHz synthetic aperture radar (SAR) tested for short-range (sub-meter) imaging. Our tested consists of a monostatic radar with synchronized transmitter and receiver, with lateral motion (over 10-30 cm) providing the SAR geometry, and range resolution provided by stepped frequency continuous wave (SFCW) signals covering a band exceeding 6 GHz. Sub-centimeter (cm) level resolution is achieved in both cross-range and slant-range. For stationary targets, platform movement provides a means of emulating a system with multiple monostatic transceivers, and exploring the impact of the spatial distribution of such transceivers. In particular, sparse spatial sampling causes significant deterioration in cross-range resolution under classical SAR image reconstruction, which is based on a point scatterer signal model. We show that performance can be greatly improved by modeling objects using “patches,” where a patch is defined as a continuum of point scatterers with (approximately) the same reflectivity. Estimation-theoretic quantities such as the Cramer-Rao Lower Bound (CRLB) and normalized cross-correlation between responses for nearby targets, are derived for both point- and patch-based signal models.

I. INTRODUCTION

Active short-range mm-wave systems have a wide range of applications including medical imaging, security scanning, autonomous vehicle navigation, and human gesture recognition. Recent advancements in silicon technology offer the possibility of realizing low-cost and highly integrated imaging systems in mm-wave (between 30 and 300 GHz) and beyond [1]. Radar imagers can be used as an alternative or complementary sensor to optical imagers. One of the main advantages of radar imagers is direct accessibility to range and Doppler information, which can be used in 3D image reconstruction and multi-target tracking applications. Robustness against weather and lighting conditions and the possibility of electronic beam-steering are other attractive properties of radar imagers.

There are various demonstrations of mm-wave imagers for different applications. For instance, [2] presents a multi-static, broadband stepped frequency continuous wave (SFCW) radar architecture for concealed weapon detection at airport security checkpoints. Other demonstrations targeting the same application are [3], [4], [5]. Another application is the use of mm-wave radar as a human gesture interface. A pulsed-radar transceiver is demonstrated in [6].

Our goal in this paper is to explore the limits of mm-wave radar when the radar transceivers are constrained to be located on a surface with relatively small form factor (e.g., a smart phone or a tablet computer), and the range of interest is of the order of a meter or less. Our focus here is on understanding the fundamentals of achievable resolution in this scenario, as a function of hardware and geometry constraints, but it is worth mentioning that one possible application is in gesture recognition.

Why 60 GHz? Moving to higher frequencies provides us with the temporal and spatial degrees of freedom that we need for high resolution imaging. Increased bandwidth availability enhances range resolution by increasing the degrees of freedom in the time-frequency domain. Cross-range resolution is enhanced by the increase in the number of spatial degrees of freedom for a constrained form factor. In particular, results derived in the context of communication [7] indicate (although detailed computations remain to be performed for radar applications) that the number of spatial degrees of freedom scale quadratically with the carrier frequency (for a 2D array with fixed form factor). The 60 GHz band is therefore an attractive candidate: 7 GHz of spectrum available at a high carrier frequency on an unlicensed basis.

Monostatic architecture: In terms of hardware choices, we have considered the monostatic architecture because of the simplicity of maintaining synchronization between a colocated transmitter and receiver. A monostatic architecture also scales easily, with only loose requirements on synchronization: a generic approach to imaging is to spread such transceivers over available surfaces as constrained by cost and form factor (whether it is a handheld or a television or a car), and then fusing the information that they gather.

Overview of the results: We point out that the geometry of short-range range imaging is quite different from that of long-range imaging in which even large objects can be modeled as point scatterers. For a handheld form factor at sub-meter ranges, our experimental and theoretical results indicate that the attainable cross-range resolution is of the order of sub-centimeter using a dense array of monostatic transceivers and a bandwidth of several GHz implemented via SFCW waveforms. For a given bandwidth and form factor, we have looked at the
impact on performance of significantly reducing the number of transceivers (spreading them over the allowable form factor), thereby constraining the available degrees of freedom. The results indicate a significant degradation in resolution of the SAR image with sparse spatial sampling. It is shown that in order to combat this dimension starvation in the signal space, we need to go beyond the point scatterer signal model and look at objects as a continuum of point scatterers, which we name as patch objects. Both experimental and theoretical results indicate that patch modelling enables the sparse array to maintain resolution, at least for simple scenes which can be efficiently characterized by patch objects.

Outline: The organization of the paper goes as follows. In Section II we briefly describe our prototype and report on experimental results for imaging a simple scene consisting two adjacent metallic plates, using both dense and sparse monostatic arrays. The results show that using a patch object rather than a point scatterer as the primitive for the signal model, gives superior performance comparing to traditional SAR processing. In Section III, we compute estimation-theoretic quantities for both point and patch signal models, that shed light on achievable resolution. Concluding remarks and future work are included in Section IV.

II. EXPERIMENTAL RESULTS

Our hardware testbed is an SFCW radar in the 54 GHz to 60.75 GHz band, built using Nonlinear Transmission Lines (NLTL) and dual high-gain horn antennas [8]. The transceiver moves in the direction parallel to the imaged scene in (1 mm) steps. The travel distance of the imager is of the order of the form factor of a portable handheld device (∼15 cm). This could emulate a number of monostatic transceivers placed on such a device. Of course, the number used here (corresponding to the number of locations, 150) is too large, but our goal is to determine first the effect of device form factor on resolution, and then to explore the effect of the sparser spatial sampling that might be used in practice. At each step of movement, complex reflection coefficients corresponding to a set of \( N = 150 \) frequencies spaced \( \Delta f = 45 \text{ MHz} \) apart, are measured and stored in a vector \( C_x \in \mathbb{C}^N \), where \( \mathbb{C}^N \) is the set of \( N \)-dimensional complex vectors. The range resolution in SFCW radar is \( \Delta r = \frac{c}{2N\Delta f} = \frac{c}{2BW} \) (∼2 cm) and maximum unambiguous range is \( r = \frac{c}{3\Delta f} \) (∼3 m), where \( c \) is the speed of light and \( BW \) denotes the total bandwidth [11].

The front end processing involves two-stage matched filtering. Using the received samples at a given transceiver location, matched filtering with respect to the range can be accomplished using the Inverse Fourier Transform. After mapping the vector of reflection coefficients to its equivalent complex range profile in this manner, we perform “SAR focusing,” or matched filtering against the expected response to each possible scatterer as a function of transceiver location.

Figure 1 shows the results of this procedure when imaging two closely spaced metallic objects (0.5 cm gap), with and without SAR focusing. Clearly, the latter is critical. If we only rely on the directionality of the horn antennas at a given location, we obtain a lateral resolution of several cm [8]. SAR focusing improves this to less than 1 cm.

In practice, we have strict limitations in terms of cost and complexity of the system, hence we need to resort to much sparser arrays than the 150 elements emulated in the preceding experiment. Figure 2 shows the result of SAR processing for the same scene of metallic objects, when deploying only 15 equispaced elements of the array. It is evident that this procedure fails to improve the cross-range resolution, and we are not able to identify the metal plates and the small gap between them. Intuitively, this is because SAR processing models each target as a point scatterer, but this model breaks down in short- and medium-range applications (e.g., gesture recognition using millimeter wave radar implemented on a handheld, or vehicular radar imaging in crowded city environments), where an object is more accurately modeled as a continuum of point scatterers. For a sparse array, there are simply not enough spatial degrees of freedom to combat the “self-interference” in the responses to this continuum of scatterers.

In view of the limitations of conventional SAR, we pro-
pose an alternative approach based on a more sophisticated primitive, in order to reduce the effective number of targets to be imaged below the dimension of the radar signal space, which is limited by parameters such as bandwidth and form factor. Specifically, we develop a patch model for a continuum of point scatterers, assuming that the patch size is small enough to approximate the reflectivity as constant across all constituent scatterers. This parsimonious model signal model, with appropriately designed estimation algorithms, allows us to “super-resolve” beyond the limits of conventional radar theory.

For example, the simple scene of metallic objects in our experiment, can be described by only two patch objects. Note that each patch is completely characterized by two parameters: width and center. Therefore, one direct consequence of patch modeling is that the description of the scene will be “sparse” in the dictionary of patch responses for the desired parameter space. Such sparse explanation gives us the opportunity of leveraging ideas from compressed sensing (CS) to recover the patch objects nested in the scene. Figure 3 shows the output of a sequential patch detection algorithm inspired by Orthogonal Matching Pursuit algorithm [10]. The description of the algorithm has been omitted due to lack of space and will appear in future publications. We see that the algorithm is able to perfectly reconstruct the scene after a few iterations using 15 array elements. Next, we investigate estimation theoretic properties of both point-scatterer and patch object signal models.

We note that there is significant recent interest in applying compressive techniques to radar using the conventional point scatterer model, with the goal of sampling at the information rate rather than at the Nyquist rate [9]. This is consistent in philosophy with our approach of using model-based estimation to make better use of the available dimensions, but we develop target primitives that go beyond the point scatterer model.

III. ESTIMATION-THEORETIC PROPERTIES

A. Point-scatterer model

Consider the geometry depicted in Figure 4. In the results reported here, we restrict attention to exploring how the SAR geometry (or equivalently, the device form factor) impacts cross-range resolution, hence we consider CW rather than SFCW signaling: even at the shorter ranges we consider, range estimates based on using larger bandwidth carry relatively little cross-range information. The signature corresponding to the point object in Figure 4 is given by

$$\mu_i(x_A) = e^{-j4\pi f_c d_{iA}/c}, \quad i \in \{1, 2, \ldots, K\}$$

where $d_{iA}$ is the distance from the object to the $i^{th}$ elements of the array. We set the model parameters to $D = 30$ cm, $K = 150$ and 15, and the length of the array aperture $L = 15$ cm.

Using this model, we compute the CRLB (assuming Gaussian noise, white across frequencies and transceiver locations) for estimating the location of two point-objects, located at $x_A$ and $x_B$, and at a fixed distance $D$ from the array. Corresponding to each object, we assume an unknown phase change introduced to the reflected signal, which is independent of the angle of incidence. Therefore, the vector of unknown parameters is extended to include point-object locations and their corresponding phase changes. It can be shown that the

![Fig. 2: Experimental results for a sparse array (15 elements) from imaging two closely spaced metallic objects: SAR processing.](image)

![Fig. 3: Experimental results for a sparse array (15 elements) from imaging two closely spaced metallic objects: Patch detection algorithm.](image)

![Fig. 4: Model of mono-static array with point-scatterer A at position $x_A$](image)
Fisher information matrix for this scenario is a function of $x_A$, $x_B$ and the difference between phase changes introduced by the two point-objects, denoted by $\phi$. We have calculated the square root of CRLB for estimating $x_A$ from the observations, where signal to noise ratio (SNR) is 10 dB, $\phi = \pi/4$ and CRLB values are normalized by $\lambda^2$. Figure 5(a) shows about a centimeter resolution for distinguishing between the two points for an array of 150 elements. Sparsifying the array by a factor of 10 introduces ambiguities in estimating $x_A$, even when $x_B$ is not in the vicinity of it, as shown in Figure 5(b). This ambiguity is a consequence of under-sampling in space.

We also consider the normalized correlation between the signature corresponding to a point $x_A$ against those corresponding to other points along a line parallel to the array, and at a distance $D$ from it. The transmitted signal is SFCW with the same parameters as in the experimental setup. Figure 6(a) shows results for $x_A = 0$ cm and $K = 150$. The size of the main lobe again indicates a resolution of about a cm for both omnidirectional and directional (main beam $\approx 20$ degrees) antennas. Figure 6(b) shows that for $K = 15$, the response of the point scatterer at $x_A = 0$ becomes highly correlated with the signatures of point scatterers far away from the center of the array. As we shall see, this leads to ambiguities for extended objects comprising infinitely many point scatterers. Note that the effective decrease in the number of degrees of freedom for directional receive antennas results in significant deterioration of lateral resolution as we move away from the center of the array. The reason is that the point scatterers close to the edge of the array can be seen by only a small fraction of the directional antennas.

![Normalized Correlation](image)

**Fig. 5:** Two slices of the CRLB for estimating $x_A$, when $x_A = 2$ cm is fixed.

**Fig. 6:** Normalized correlations for a point scatterer at $x_A = 0$

### B. Patch-based model

We first introduce the patch model, and then compute associated estimation-theoretic quantities such as the CRLB and normalized correlations between noiseless received signals. Consider the geometry depicted in Figure 7. We consider a one-dimensional patch uniquely identified by its starting and ending points, denoted by $x_A$ and $x_B$, respectively. Alternatively, we can use centre $O \triangleq (x_A+x_B)/2$ and width $W \triangleq x_B-x_A$ to identify the same patch. A key approximation inherent in this primitive is that the reflectivity is constant across the patch (thus, we must choose our nominal patch size small enough that this approximation is valid). Similar to the analysis for point scatterer, we consider CW rather than SFCW for CRLB analysis. The response at $i^{th}$ element of the array is given by

$$s_i(x_A, x_B) = \int_{x_A}^{x_B} e^{-j2\pi f_{c} d_{i}\cos \theta} d\theta.$$

The received signal power at $i^{th}$ element is $P_i = |s_i(x_A, x_B)|^2$, and the total received power $P = \sum_{i=1}^{K} P_i$. The $K \times 1$ array response is modeled as

$$y = s(x_A, x_B) + z,$$

where $z \sim CN(0, 2\sigma^2 I_K)$ denotes additive complex Gaussian noise. The Fisher information matrix for estimating $x_A$ and $x_B$...
$x_B$ is given by

$$F(x_A, x_B) = \frac{1}{\sigma^2} \left[ \begin{array}{cc} K & -\sum_i \cos \left( \frac{4\pi}{\lambda}(d_{iA} - d_{iB}) \right) \\ -\sum_i \cos \left( \frac{4\pi}{\lambda}(d_{iA} - d_{iB}) \right) & K \end{array} \right].$$

We can use $F(x_A, x_B)$ to find CRLB for the variance of any unbiased estimator of a linear combination of $x_A$ and $x_B$ using the following well-known fact: for $a \in \mathbb{R}^n$, the CRLB for $a^T \theta$ is given by $a^T F^{-1}(\theta) a$, where $F(\theta)$ is the Fisher information matrix for $\alpha$-dimensional parameter vector $\theta$. We use this to calculate the CRLB for estimation of the width $W$ and center $O$ for a patch. We set the model parameters to $D = 30$ cm, $K = 150$ and 15, and length of the array $L = 15$ cm. Since we are interested in the effect of form factor on the estimation bounds, we fix the SNR = 10 dB for the entire parameter space. Figure 8 shows the square root of normalized CRLB for estimating $W$. We see that sparsifying the array causes ambiguity in only a small portion of the parameter space (similar observations hold for the CRLB for estimating $O$). To better see how these bounds compare to one another, we fix the center at $O = 0$ cm, and calculate CRLB for different values of $W$. Figure 9 shows that CRLB for $K = 15$, closely follows the bound for $K = 150$. Therefore, for a fixed SNR, sparsifying the array does not lead to fundamental ambiguities for estimating the parameters of a patch object within a large fraction of the parameter space.

We also look at normalized correlations and compare the results for point- and patch-based signal models. To this end, let us construct two inventories of responses: (1) the inventory of signatures corresponding to point scatterers along a line parallel to the array, and at a distance $D = 30$ cm from it; (2) the inventory of signatures corresponding to patch objects having a nominal width $W_{\text{nom}} = 1$ cm, along the same line parallel to the array. The transmitted signal is SFCW with the same parameter as the experimental setup.
Figure 10 shows the result of the correlations of a patch response whose width is exactly equal to $W_{\text{nom}}$, against the two inventories of responses that we have constructed. The size of the main lobe again indicates a resolution of about a cm for both inventories. However, the mainlobe to sidelobe ratio for the patch inventory is much larger (i.e., better) than that for the point inventory.

![Diagram showing normalized correlations for wide and interfering patches](image)

(a) Wide patch  
(b) Interfering patches

Fig. 11: Normalized correlations for (a) a wide patch with $W > W_{\text{nom}}$, and (b) multiple patches

The advantage of the patch-based model becomes more obvious when the scene contains wider patches, or multiple patches interfering with each other. Figure 11(a) show the result of correlations for a patch with $W = 5$ cm. We see that correlations against the patch inventory are able to describe the scene very well, even though $W > W_{\text{nom}}$, whereas correlations against point signatures fails to provide an unambiguous explanation of the scene. Figure 11(b) shows the result of correlations for a scene consisting multiple patches with different widths. We see that correlations against patch responses is much more robust against destructive interference between signals from multiple objects. As shown in Section II, experimental results also indicate that correlations against patch objects provide the required information to detect patches, and to reconstruct simple scenes unambiguously.

IV. CONCLUSION

We have presented experimental and theoretical results showing that it should be possible to obtain sub-cm resolution using mm wave radar implemented in handheld-sized form factors, thus opening up interesting potential applications such as gesture recognition. However, standard SAR processing, which is based on the classical point scatterer target model, deteriorates for sparse spatial sampling at short and medium ranges, because each object appears as a continuum of point scatterers which are difficult to resolve with a limited number of spatial degrees of freedom. We show, however, that our proposed new patch-based primitive is effective in providing a sparse description of the scene. Super-resolution algorithms based on this primitive yield improved image reconstruction on our testbed, and estimation-theoretic computations provide further support for the effectiveness of our approach.

Ongoing research includes incorporation of Doppler information for moving targets, extensive validation for more complex scenes in the context of specific applications, waveform design, performance/complexity comparisons with multistatic architectures, and low-power hardware design.

V. ACKNOWLEDGMENT

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.

REFERENCES


