

SAR Canonical Feature Extraction using Molecule Dictionaries

G. Barry Hammond and Julie Ann Jackson
 Department of Electrical and Computer Engineering
 Air Force Institute of Technology
 Wright-Patterson AFB, OH 45433
 Email: julie.jackson@afit.edu

Abstract—We apply a molecule dictionary approach to synthetic aperture radar canonical feature extraction. These canonical features capture physically-relevant scattering geometry as a function of shape type, frequency, aspect, and polarization. The extraction problem is a nonlinear nonconvex optimization that includes model order selection, feature classification, and parameter estimation. Previous work used image-based initializations, gradient descent, and a hierarchical classification scheme to extract the features. The dictionary approach shifts much of the computational burden to dictionary formation which can be done offline, prior to feature extraction. We show results for cases when the true feature lies in the dictionary and when it does not. Discussion of the practical challenges of dictionary construction is given in the context of recent sparse recovery literature.

I. INTRODUCTION

Attributed scattering feature models have shown potential in aiding automatic target recognition and scene visualization from radar scattering measurements [1]–[6]. An example processing flow is shown in Figure 1. Attributed scattering features capture physical scattering geometry, including the limited response of target scattering over wide angles, that is not discerned from traditional point scatter models. For example, the electromagnetic target signature in Figure 1 may be well-represented by the depicted set of canonical shapes. However, the increased fidelity of attributed models also increases the difficulty of extracting such features from measurements [4], [5]. Efficient and accurate extraction of physically-meaningful features from synthetic aperture radar (SAR) remains a challenging problem. In this paper, we apply a molecule dictionary approach, similar to methods in [7]–[9], to the problem of extracting the canonical features of [6] (plates, cylinders, etc.) from SAR phase history data.

The proposed approach evaluates correlation between dictionary entries, referred to as *atoms*, and clusters similar atoms into *molecules*. By first comparing the molecules to the data, then extracting the appropriate atoms, the computational load of the dictionary approach is greatly reduced. A previous approach to extracting the canonical features of [6] used an image-based initialization scheme in conjunction with gradient descent parameter estimation and a hierarchical shape classifier [5]. Application of the molecule method to the particular scattering models in [6] shifts the bulk of the computational burden from the feature extraction (gradient descent) algorithm to dictionary formation which can be

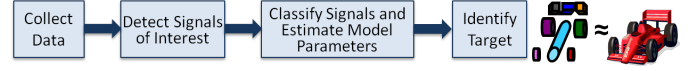


Fig. 1. Data processing chain and cartoon of target identification from canonical shape features extracted from the target’s electromagnetic signature.

done offline. We are motivated to overcome the correlation limitations of the canonical feature dictionary because the geometric objects (plates, dihedrals, cylinders, etc.) provide more easily visualized representations of the scene than do alternative basis dictionaries such as wavefronts or chirps [10] or point scatterers [7].

We begin the paper by presenting the molecule dictionary formulation and feature extraction algorithm. Although the methods are straightforward, several challenges arise: how coarsely the dictionary is sampled, dictionary structure, computer memory requirements, and estimator accuracy/confidence. Rather than provide an extensive literature review up front, we discuss specific challenges in relation to the literature in Section V, after presenting initial results. The feature extraction application exemplifies the practical challenges of achieving a physically-relevant sparse reconstruction for high-dimensional signals.

II. FEATURE EXTRACTION PROBLEM

For a noisy set of radar phase history measurements $\mathbf{Y}(\Lambda)$, the feature extraction problem may be posed as the least-squares problem

$$\min_{\Theta, \Gamma} \left\| \mathbf{Y}(\Lambda) - \sum_{m=1}^{\hat{M}} \mathbf{S}_{\Gamma_m}(\Lambda; \Theta_m(\Gamma_m)) \right\|_2^2 \quad (1)$$

for radar parameters Λ , feature types $\Gamma = [\Gamma_1, \dots, \Gamma_{\hat{M}}]$ and feature parameters $\Theta = [\Theta_1, \dots, \Theta_{\hat{M}}]$. We use the shorthand notation Λ to represent operating parameters of the radar. In this case, we consider a monostatic radar whose parameters include: wavenumber $\mathbf{k} \triangleq \frac{2\pi\mathbf{f}}{c}$, P polarization channels, and a flight path parameterized by azimuth (ϕ) and elevation (θ) aspects. The functional form of $\mathbf{S}_{\Gamma_m}(\Lambda; \Theta_m(\Gamma_m))$ depends on the scatterer type indicated by Γ_m and the corresponding parameter set Θ_m . The vector Θ_m contains physical parameters such as size, location, and orientation corresponding to the m th scatterer, as described in [6]. The optimization in (1)

represents the feature extraction problem for any scattering model \mathbf{S} . Aspect-dependent point scatterers and wave-based dictionaries have been considered in previous work [7], [10]. We consider the six canonical scattering types in [6] such that $\Gamma_m \in \{\text{plate, dihedral, trihedral, cylinder, top-hat, sphere}\}$.

Suppose that these six canonical scatterers, parameterized by their location, size, and orientation, represent all possible radar scattering responses. Sampling the parameter space results in a discrete dictionary \mathcal{D} of possible radar scattering responses. Sampling the observation space (radar operating parameters Λ) may be written as a matrix H that extracts the observed rows of \mathcal{D} . Vector ψ is a set of coefficients that combines the dictionary terms to form the measured data as

$$\mathbf{Y}(\Lambda) = H\mathcal{D}\psi + \mathbf{W} \quad (2)$$

for i.i.d. Gaussian noise \mathbf{W} . Since $\mathbf{D} = H\mathcal{D}$ is the effective dictionary for a given measurement scenario, we consider \mathbf{D} in place of \mathcal{D} in the following. Dictionary \mathbf{D} may be split into sub-dictionaries \mathbf{D}_{Γ_j} for shape types $j = 1 \dots J$. Each sub-dictionary takes the form

$$\mathbf{D}_{\Gamma_j} = [\mathbf{d}(\Lambda; \Theta_1(\Gamma_j)) \cdots \mathbf{d}(\Lambda; \Theta_{N_j}(\Gamma_j))] \quad (3)$$

with N_j columns, referred to as *atoms*, that sample the parameter space $\Theta(\Gamma_j)$. In (3) each column is a $2K \times 1$ vector

$$\mathbf{d}(\Lambda; \Theta_n(\Gamma_j)) = \begin{bmatrix} \text{Re} \{ \mathbf{S}_{\Gamma_j}(\Lambda; \Theta_n(\Gamma_j)) \} \\ \text{Im} \{ \mathbf{S}_{\Gamma_j}(\Lambda; \Theta_n(\Gamma_j)) \} \end{bmatrix} \quad (4)$$

that samples the radar phase history response along operating parameters Λ .

The problem in (1) includes model order selection, feature classification, and parameter estimation. Assuming model order $M \ll N_a$ total dictionary atoms, we can recast (1) as a sparse signal recovery problem [11]–[13]

$$\min \|\psi\|_{\ell_0} \quad \text{s.t.} \quad \|\mathbf{Y}(\Lambda) - \mathbf{D}\psi\|_2^2 < \epsilon \quad (5)$$

for coefficients $\psi = [\psi_1, \dots, \psi_{N_a}]^T$. Then, the ℓ_0 -norm of the resulting ψ is the model order estimate, and shape class and parameter estimates are given by the selected dictionary entries. However, solution of (5) is of combinatoric complexity. Relaxing the sparsity constraint to the ℓ_1 norm, (5) may be formulated as the convex basis pursuit denoising (BPDN) problem [12]

$$\min \|\mathbf{Y}(\Lambda) - \mathbf{D}\psi\|_2^2 + \lambda \|\psi\|_1 \quad (6)$$

for a given level of sparsity λ and dictionary \mathbf{D} .

Although BPDN is convex, its success depends on the coherence properties of the dictionary. A dictionary that satisfies the restricted isometry property (RIP) has provable convergence to ‘good’ solutions [11]. Dictionaries populated with random elements satisfy the RIP with high probability, but general dictionaries do not. Mutual coherence between atoms violates RIP, and BPDN reconstruction performance is not guaranteed. For the canonical shape dictionaries, high inter-atom correlation occurs for fine sampling of the parameter space and for limited diversity in the observation space.

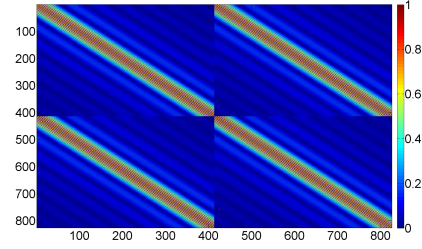


Fig. 2. Magnitude of Gram matrix for a dictionary comprised of dihedral and plate samples for a linear flight path and VV polarization

Confusion may arise between different parameterizations of the same shape (e.g. length of 1m versus length of 0.99m) or between shape class when limited data is observed. For example, a constant elevation flight path that captures VV polarized specular reflections from a plate or a dihedral of the same length (at the same location) results in identical radar phase history data [14]. Figure 2 shows the magnitude of the Gram matrix $\mathbf{G} = \mathbf{D}^H \mathbf{D}$ for this example of dihedral-plate correlation. Correlation near the diagonal bands is from similar parameter sets. The upper and lower diagonal bands are correlations between similar dipoles and plates. Clustering of correlated atoms into *molecules* allows for a reduced dictionary comprised of representative elements, such as mean vectors in a k-means clustering. The reduced molecule dictionary should have low inter-column correlation, resulting in a diagonal Gram matrix, and thus improve the BPDN convergence [7]–[9]. Computational speed also improves as the size of the dictionary matrix is reduced.

III. ALGORITHM

We apply the molecule dictionary approach to canonical shape feature extraction. The proposed algorithm estimates one scatterer at a time using an iterative model and subtract approach similar to that in [5] but replaces image formation, initialization, and gradient descent optimization with BPDN applied to molecules. The proposed algorithm is an extension of the molecule-based approach to extracting aspect-dependent point scatterers in [7], the general tree-based pursuit algorithm in [9], and the music signal decomposition in [8]. We refer to both the representative vectors and the set of atoms it represents as molecules; context should make usage clear.

The molecule dictionary approach to canonical shape feature extraction is as follows.

- 1) Form dictionaries \mathbf{D}_j for each canonical shape.
- 2) Reduce each dictionary to molecule representation $\bar{\mathbf{D}}_j$ by clustering atoms until a user-defined maximum coherence threshold is satisfied: $\max |\bar{\mathbf{D}}_j^H \bar{\mathbf{D}}_j| < \eta$.
- 3) Collect measured phase history data.
- 4) Run BPDN on all reduced dictionaries to solve $\min_{\psi_j} \|\mathbf{Y}(\Lambda) - \bar{\mathbf{D}}_j \psi_j\|_2^2 + \lambda \|\psi_j\|_1$.
- 5) Choose the molecule BPDN result that best balances sparsity with largest coefficients according to $\max_j (\max_i |\psi_j(i)| - \lambda \|\psi_j\|_1)$.

TABLE I
CANONICAL SHAPE PARAMETER VALUES, LISTED AS A SCALAR VALUE OR MINIMUM:STEP-SIZE:MAXIMUM FOR VARIED PARAMETERS.

Shape	Location			Size			Orientation		
	x (m)	y (m)	z (m)	Height H (m)	Length L (m)	Radius r (m)	Roll $\tilde{\gamma}^\circ$	Pitch $\tilde{\theta}^\circ$	Yaw $\tilde{\phi}^\circ$
plate	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	0.5 : 0.5 : 3	0.5 : 0.5 : 3	-	0	-30	0
dihedral	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	0.5 : 0.5 : 3	0.5 : 0.5 : 3	-	0	0	0
trihedral	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	0.5 : 0.5 : 3	-	-	0	0	0
top-hat	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	0.5 : 0.5 : 3	-	0.5 : 0.5 : 3	0	0	0
cylinder	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	-	0.5 : 0.5 : 3	0.5 : 0.5 : 3	0	0	0
sphere	-10 : 5 : 10	-10 : 5 : 10	-10 : 5 : 10	-	-	0.5 : 0.5 : 3	-	-	-

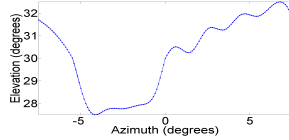


Fig. 3. Radar flight path used in simulation results.

- 6) Compute the posterior log-likelihood for each atom in the selected molecule and estimate the true atom as the maximizer (MAP=LS for uniform priors).
- 7) Subtract the phase history corresponding to the selected atom from measured data.
- 8) Return to Step 4 until a desired amount of energy is modeled.

Steps 1 and/or 4 may be augmented by SAR image formation, polarimetric decomposition, and other processes to pre-screen shape class and initialize parameter estimates (as in [5]) to limit the number and extent of dictionaries searched.

IV. RESULTS

The algorithm presented in Section III is tested in [15] for monostatic SAR over a variety of radar and object scenarios. Here, we present a summary of results for a relatively diverse set of radar observations: 151 uniformly-spaced samples along the non-constant elevation path defined in [14] and shown in Figure 3; 125 frequency samples uniformly-spaced over a bandwidth of 300MHz centered at 10GHz; fully-polarimetric, e.g. HH, HV, and VV channels. SAR data is simulated from the feature models in [6] to test the algorithm performance without model mismatch to observations. The parameter values for the dictionary are listed in Table I as min:step:max values. Coarse dictionary sampling and known object orientations are used due to memory constraints, discussed in Section V-A.

Three scenarios are tested. Scenario 1 extracts a single shape, whose truth corresponds to an atom in the original dictionary, observed over the radar parameters listed above with peak signal-to-noise ratio (SNR) 30dB. Molecule dictionaries were formed with maximum coherence $\eta = 0.2$. BPDN solutions for each shape's molecule dictionary and sparsity constraint $\lambda = 0.2$ are shown in Figure 4. The correct molecule, dihedral molecule 33, is selected by the criteria in algorithm Step 5. Subsequently, the 60th atom in molecule 33 is correctly selected from the posterior least-squares estimator. The results in Figure 4 are for a single realization of the noisy signal observation. Figure 5 shows composite results

for estimating each shape type under Scenario 1 as a function of SNR and RCS. Each SNR result is the percentage correct over 100 Monte Carlo trials.

Scenario 2 extends Scenario 1 to test the iterative model and subtract scheme for extracting multiple objects whose parameters are included in the dictionary atoms. For a scene comprised of a plate with $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (10\text{m}, 0\text{m}, 0\text{m}, 1\text{m}, 1\text{m}, 0^\circ, -30^\circ, 0^\circ)$, a dihedral with $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (0\text{m}, 10\text{m}, 5\text{m}, 1\text{m}, 2\text{m}, 0^\circ, -30^\circ, 0^\circ)$, and top-hat with $(x, y, z, H, r, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (-10\text{m}, 5\text{m}, 0\text{m}, 3\text{m}, 1\text{m}, 0^\circ, 0^\circ, 0^\circ)$, the algorithm extracts all objects correctly. Maximum coefficients for each shape dictionary, sparsity, and selection criteria are listed in Table II for three iterations. The shape selected at each iteration is highlighted—the correct shape (with correct parameter estimates) is selected.

TABLE II
MAXIMUM COEFFICIENTS FROM EACH SHAPE DICTIONARY ALONG WITH THE ℓ_1 NORM FOR SCENARIO 2; ROW OF SELECTED SHAPE AT EACH ITERATION IS HIGHLIGHTED.

Shape	Iteration	$\max \psi $	$\ \psi\ _1$	$\max \psi - \lambda \ \psi\ _1$
plate	1	0.4396	0.6037	0.3189
	2	0.4396	0.6037	0.3189
	3	0.4396	0.6037	0.3189
dihedral	1	0.8149	0.9083	0.6333
	2	0.0076	0.0651	-0.0054
	3	0.0024	0.0352	-0.0046
sphere	1	0.4269	24.1672	-4.4065
	2	0.4269	24.1672	-4.4065
	3	0.4269	24.1672	-4.4065
top-hat	1	1.7148	6.1812	0.4786
	2	1.7140	2.9059	1.1329
	3	0.0108	1.1921	-0.2276
trihedral	1	3.8552×10^{-4}	0.0255	-0.0047
	2	3.8552×10^{-4}	0.0255	-0.0047
	3	3.8552×10^{-4}	0.0255	-0.0047
cylinder	1	0.8770	10.2192	-1.1668
	2	0.8770	10.2192	-1.1668
	3	0.8770	10.2192	-1.1668

Finally, Scenario 3 attempts to extract a single shape, where the true parameter set corresponds to a point between dictionary samples, i.e. an atom that is NOT in the dictionary. Flightpath, SNR, λ , and η values are as in Scenario 1. The true shape is a plate with parameters $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (4\text{m}, 0\text{m}, 0\text{m}, 1\text{m}, 2\text{m}, 0^\circ, -30^\circ, 0^\circ)$. The nearest dictionary sample is at $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (5\text{m}, 0\text{m}, 0\text{m}, 1\text{m}, 2\text{m}, 0^\circ, -30^\circ, 0^\circ)$, which is clustered in plate molecule 24. However, Figure 6 shows that the BPDN coefficient estimated for molecule 24 is near zero; the selected largest molecule

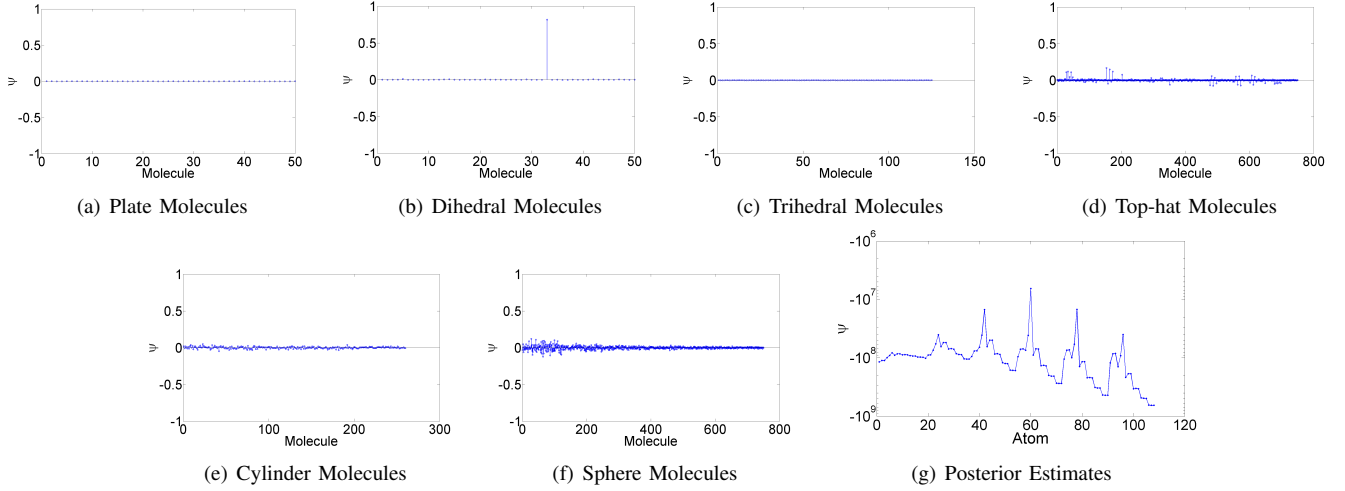


Fig. 4. (a-f) Scenario 1 BPDN results for canonical shape molecule dictionaries when true shape is a dihedral with parameters $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (0\text{m}, 10\text{m}, 5\text{m}, 1\text{m}, 2\text{m}, 0^\circ, 0^\circ, 0^\circ)$, corresponding to dihedral molecule 33. (g) Posterior estimates for each atom within selected dihedral molecule (max at true atom = 60).

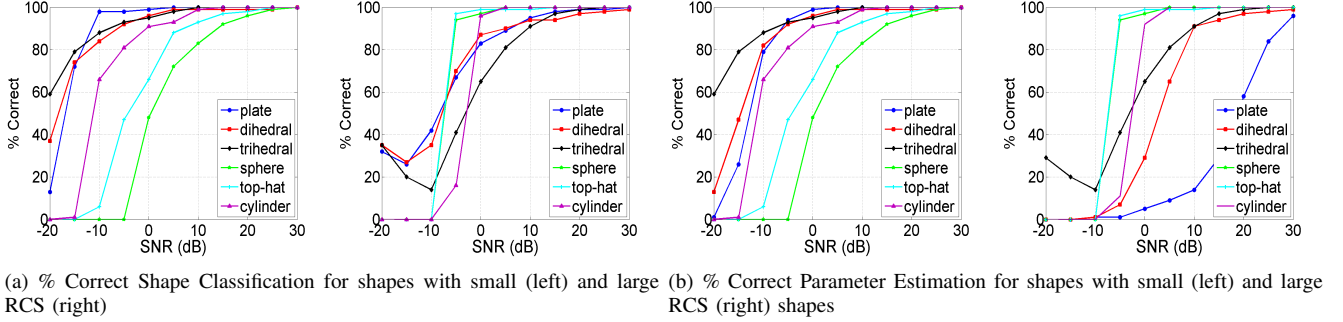


Fig. 5. Percentage of correct shape classification and correct atom parameter set selection for small 0.5m-3m size and large 6.5m-9m size targets.

coefficient is plate molecule 10. The extracted atom from molecule 10 has parameters $(x, y, z, L, H, \tilde{\gamma}, \tilde{\theta}, \tilde{\phi}) = (10\text{m}, 5\text{m}, -10\text{m}, 0.5\text{m}, 3\text{m}, 0^\circ, -30^\circ, 0^\circ)$. This extracted atom far from the true plate location is a consequence of the coarse dictionary sampling. Figure 6 shows the range profile for the extracted plate is a better match to the true range profile than that for the nearest dictionary sample. Although the estimated atom is the best representative signal, the underlying physical parameters are incorrect, so the solution is not desirable. More discussion on appropriate dictionary sampling is given in Section V-C.

V. DISCUSSION

Initial results are promising, but several practical challenges impacting successful feature extraction are discussed next.

A. Memory requirements

Generating the required feature dictionaries requires large amounts of computer memory, even for coarsely-sampled feature parameters. Each canonical shape response is represented by four to eight size, location, and orientation parameters in Θ_j , depending on the shape type. The dictionary \mathbf{D} contains $N_a = \sum_{j=1}^J \prod_{q=1}^{Q_j} L_{jq}$ atoms, corresponding to L_{jq}

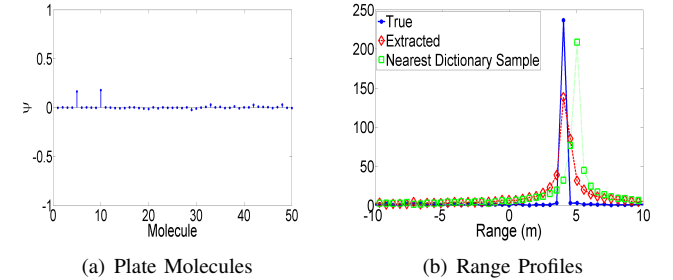


Fig. 6. (a) Scenario 3 BPDN results for the selected plate molecule dictionary when true shape lies between dictionary samples. (b) Range profile comparison of true, extracted, and nearest dictionary sample.

samples of the $q = 1, \dots, Q_j$ parameters for the j th shape. In addition, each atom has $2K$ entries, including both real and complex values of radar returns associated with multiple frequencies, aspects, and polarizations. Assuming a 64-bit precision machine, the memory required to store a dictionary for K observations is $8 \times 2KN_a$ bytes. Even for the short flight path and coarsely-sampled dictionary in Table I, the memory requirement is 16.5GB for $N_a = 19500$ atoms. Fine sampling of parameters is typically desired to avoid quantiza-

tion error and achieve accurate representations of the observed scene. Clustering similar atoms into molecules reduces the memory requirements and gives us the desired dictionary coherence properties for successful BPDN. An extension of the demonstrated approach for very large dictionaries is to perform molecule clustering over portions of dictionaries that can be held in memory at one time. Then successive clustering can group similar molecules to build up a hierarchical search method as in [7], [9].

B. Dictionary Structure

In general, \mathbf{D} is a non-orthogonal dictionary. The atoms do not form a basis set but are assumed to span the space of all possible scatterers (for a given set of radar parameters). The convergence properties in [11] apply to dictionaries whos entries form an over-complete basis set. Fortunately, the properties extend to the general case as shown in [13]. Future work will consider properties and construction of (molecule) dictionaries for guaranteed convergence under [13]. Moreover, we seek convergence to a solution that lies in the subset of sparse solutions that includes only those canonical shapes that provide a physically-relevant representation of the scene for human-aided target recognition and/or scene visualization.

C. Dictionary Sampling

Scenario 3 demonstrated that coarse dictionary sampling can result in quantization error that does not give the desired result. This is expected, but how fine should one sample? In [16], the relation between sampling resolution and root mean-squared error (RMSE) is studied for a decaying exponential model. Results in [16] indicate that sampling too coarsely or too finely can negatively impact RMSE. We explore the sampling issue for our dictionary with a simple example. Consider the problem of estimating the x location (down-range) parameter of the plate in Scenario 3. For this example, we consider two locations $x = 0\text{m}$ and $x = 0.25\text{m}$.

Figure 7 shows the least-squares (LS) cost function with fine and coarse sampling for true values of x that lie both in and out of the cost surface (dictionary) sample grid. The envelope of the LS cost corresponds to the non-coherent (NC) cost function utilized in [4], [5] and is also plotted. For fine sampling, local minima in the LS cost function are apparent. The LS global minima occurs on the true x location when it is a point in the sample grid or at the nearest sample $x = 0.2498\text{m}$ when the true point $x = 0.25\text{m}$ is not in the sample grid. However, very fine sampling is computationally costly. One might be motivated to sample the cost function at the radar range resolution, since any sub-pixel parameter estimates would be considered a super-resolution technique. Figures 7(c) and 7(d) show the LS and NC cost functions sampled at the radar range resolution $c/(2BW) \approx 0.5\text{m}$ for the on-grid and off-grid true locations. In the on-grid case, the LS and NC criteria achieve zero cost at the true location $x = 0\text{m}$. However, in the off-grid case, the minimum LS cost occurs at $x = -0.9993\text{m}$, far from the true value. Thus, the range resolution sampling is too coarse to accurately estimate

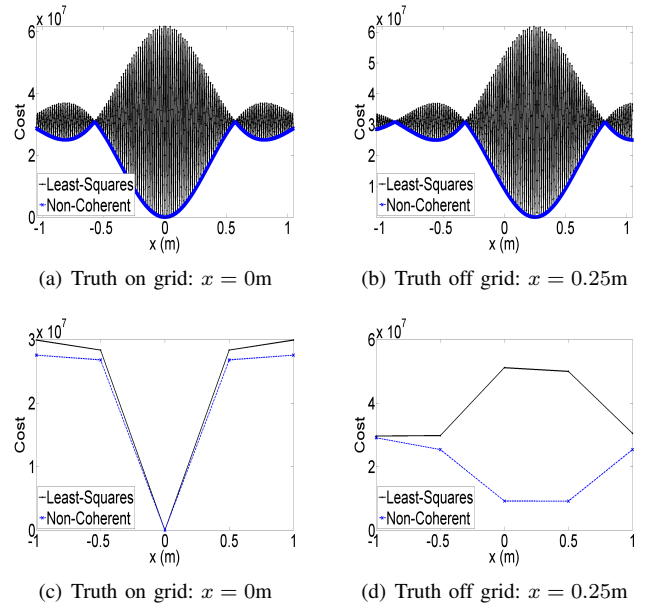


Fig. 7. Cost functions for unknown plate x location sampled at $1/300\text{th}$ range resolution $\approx 1.66 \times 10^{-3}\text{m}$ (top row) and the range resolution $\approx 0.5\text{m}$ (bottom row) for $\text{SNR} = \infty$.

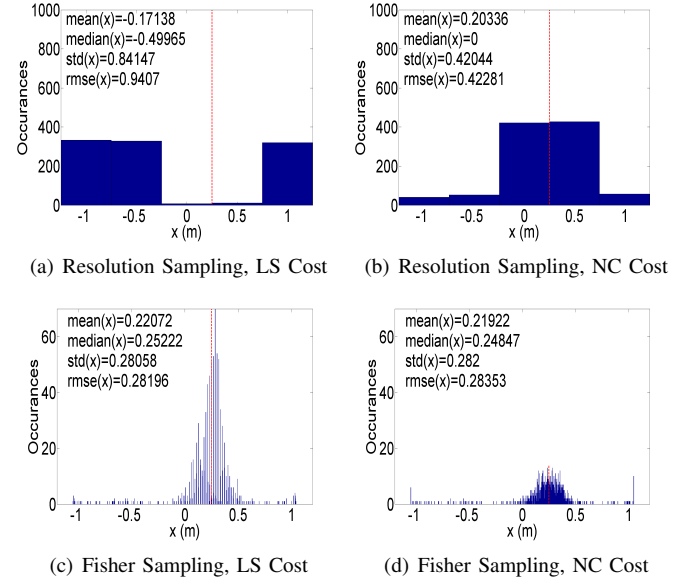


Fig. 8. Histograms of 1000 Monte Carlo trials estimating plate x parameter for $\text{SNR}=30\text{dB}$. Vertical red dashed line indicates true parameter location.

the plate location when it is not located at the center of a resolution cell, i.e. it is between samples in the dictionary.

The NC cost outperforms the LS cost in the off-dictionary case—the local minima occurs at the nearest on-dictionary sample at 0.5m for the coarse sampling case. In general, the envelope cost is more robust to initialization errors and the trappings of local minima than is the LS cost. However, careful initialization is required for both the LS and NC cost functions for numerical search algorithms to reach the global solution. In the LS case, computationally expensive search methods are almost certainly needed to locate the global rather than local

minima. Thus, the NC cost should be considered in future versions of the dictionary extraction method. However, when estimating multiple parameters, the local minima of the NC cost compound, so the advantage is not as great as it appears in the one-dimensional searches shown in Figure 7.

Expected noise levels should also factor into choice of cost function and sampling. One should consider sampling at the limit of what is estimable for a given scenario. Then, computational resources may be variably allocated, depending on the problem at hand. The Fisher information sampling rate $(\mathcal{I})^{-1/2}$ is defined in [16] to do just that. (Matrix \mathcal{I} is the Fisher Information matrix). For parameter vector Θ , the sampling rate is the square root of the Cramer-Rao bound (CRB) for each parameter. The Fisher information sampling rate says we should sample infinitesimally fine for the noiseless case, as the CRB is zero. For real-world scenarios, SNR is limited and the CRB is non-zero. For the plate scenario under consideration, the Fisher information sampling rate is $\Delta x = 2.497 \times 10^{-3}\text{m}$ for an SNR of 30dB. It is clear that for targets on the sampling grid, estimates will be grouped near the global optimum. Figure 8 shows histograms of the estimated x locations for the off-grid case for 1000 Monte Carlo realizations of the plate in noise for range resolution and Fisher information rate sampling. As expected, the Fisher sampling results in a smaller RMSE for the estimated value of x . Thus, sampling the dictionary at the Fisher rate is desirable.

The Fisher information sampling rate results in a very large dictionary when considering estimation of multiple parameters. The successive clustering approach would make such a sampling rate feasible and should be explored in future work. Alternatively, true parameter sets not captured in the dictionary sampling can be considered perturbations of the dictionary matrix [17], and application of the sparsity-cognizant total least-squares in [17] should be considered as an alternative solution method for canonical feature extraction.

D. Estimator Confidence

Proper sampling and dictionary construction should yield ‘good’ solutions to the sparse reconstruction problem [13], [16]. Nonetheless, it is desirable to have a confidence measure for each extracted feature. Bayesian estimation can provide such a measure in terms of the posterior probability. The proposed algorithm is set up to extract atoms from selected molecules by maximizing the posterior. However, results thus far have utilized uniform priors, so the maximization is equivalent to least-squares estimation. In [18], Laplacian priors on dictionary coefficients ψ are considered for the sparse reconstruction problem. Future work will explore incorporating the posterior density function calculation to allow for a true MAP estimation in algorithm Step 6. Then, in cases where the MAP estimate results in small likelihood values, the algorithm can continue but report low confidence in the estimate.

VI. CONCLUSION

We have presented an algorithm for canonical feature extraction based on the molecule-based approaches in [7]–[9]

and shown initial results for a simple, two-tier (molecules and atoms) dictionary. Practical challenges of dictionary sampling, structure, computer memory, and estimator confidence motivate further work, as discussed in Section V. Extension of the simple cases demonstrated in the literature to this more complicated problem may provide a means for robust canonical feature extraction.

ACKNOWLEDGMENT

Thanks to Mr. Jason Parker for comments on early drafts. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government. Approved for public release: #88ABW-2012-5688.

REFERENCES

- [1] L. C. Potter and R. L. Moses, “Attributed scattering centers for SAR ATR,” *IEEE Trans. on Image Processing*, vol. 6, no. 1, pp. 79–91, 1997.
- [2] M. J. Gerry et al., “A parametric model for synthetic aperture radar measurements,” *IEEE Trans. on Antennas and Propagation*, vol. 47, no. 7, pp. 1179–1188, Jul. 1999.
- [3] M. A. Koets and R. L. Moses, “Image domain feature extraction from synthetic aperture imagery,” in *IEEE Intern’l Conf. on Acoustics, Speech, and Signal Processing*, vol. 4, Mar. 1999, pp. 2319–2322.
- [4] J. A. Jackson, *Three-Dimensional Feature Models for Synthetic Aperture Radar and Experiments in Feature Extraction*. The Ohio State University: Ph.D. Dissertation, 2009.
- [5] J. A. Jackson and R. L. Moses, “Synthetic aperture radar 3D feature extraction for arbitrary flight paths,” *IEEE Trans. on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2065–2084, Jul. 2012.
- [6] J. A. Jackson, B. D. Rigling, and R. L. Moses, “Canonical scattering feature models for 3D and bistatic SAR,” *IEEE Trans. on Aerospace and Electronic Systems*, vol. 46, no. 2, pp. 525–541, Apr. 2010.
- [7] K. R. Varshney et al., “Sparse representation in structured dictionaries with application to synthetic aperture radar,” *IEEE Trans. on Signal Processing*, vol. 56, no. 8, pp. 3548–3561, Aug. 2008.
- [8] L. Daudet, “Sparse and structured decompositions of signals with the molecular matching pursuit,” *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 14, no. 5, pp. 1808–1816, Sep. 2006.
- [9] P. Jost, P. Vanderghynst, and P. Frossard, “Tree-based pursuit: Algorithm and properties,” *IEEE Trans. on Signal Processing*, vol. 54, no. 12, pp. 4685–4697, Dec. 2006.
- [10] M. R. McClure and L. Carin, “Matching pursuits with a wave-based dictionary,” *IEEE Trans. on Signal Processing*, vol. 45, no. 12, pp. 2912–2927, Dec. 1997.
- [11] E. Candes, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [12] S. S. Chen, D. Donoho, and M. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Review*, vol. 43, no. 1, pp. 129–159, 2001.
- [13] D. Donoho and M. Elad, “Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization,” in *Proc. of the National Academy of Sciences*, Mar. 2003, pp. 2197–2202.
- [14] J. A. Jackson and R. L. Moses, “Identifiability of 3D attributed scattering features from sparse nonlinear apertures,” in *Proc. of SPIE, Algorithms for Synthetic Aperture Radar Imagery XIV*, E. G. Zelnio and F. D. Garber, Eds., vol. 6568, 2007.
- [15] G. B. Hammond, *Target Classification of Canonical Scatterers Using Classical Estimation and Dictionary Based Techniques*. Air Force Institute of Technology: MS Thesis, Mar. 2012.
- [16] C. D. Austin et al., “On the relation between sparse sampling and parametric estimation,” in *Proc. IEEE 13th DSP Workshop 5th Signal Processing Ed. Workshop (DSP/SPE 2009)*, Jan. 2009, pp. 387–392.
- [17] H. Zhu, G. Leus, and G. B. Giannakis, “Sparsity-cognizant total least-squares for perturbed compressive sampling,” *IEEE Trans. on Signal Processing*, vol. 59, no. 5, pp. 2002–2016, May 2011.
- [18] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Trans. on Signal Processing*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.