Expectation-maximization approach to target model generation from multiple synthetic aperture radar images

John A. Richards
Alan S. Willsky, MEMBER SPIE
John W. Fisher III
Massachusetts Institute of Technology
Laboratory for Information and Decision Systems
77 Massachusetts Avenue
Cambridge, Massachusetts 02139
E-mail: johnrichards@alum.mit.edu

Abstract. A key issue in the development and deployment of model-based automatic target recognition (ATR) systems is the generation of target models to populate the ATR database. Model generation is typically a formidable task, often requiring detailed descriptions of targets in the form of blueprints or CAD models. We propose a method for generating a 3-D target model directly from multiple SAR images of a target obtained at arbitrary viewing angles. This 3-D model is a parameterized description of the target in terms of its component reflector primitives. We pose the model generation problem as a parametric estimation problem based on information extracted from the SAR images. We accomplish this parametric estimation in the context of data association using the expectation-maximization (EM) method. Our model generation technique operates without supervision and adaptively selects the model order. Although we develop our method in the context of a specific data extraction technique and target parameterization scheme, our underlying framework is general enough to accommodate different choices. We present results demonstrating the utility of our method. *© 2002 Society of Photo-Optical Instrumentation Engineers.* [DOI: 10.1117/1.1417493]

Subject terms: synthetic aperture radar; model-based automatic target recognition; target model generation; scattering models; data association; expectation-maximization method; parametric estimation; radar cross section.

Paper 200229 received June 8, 2000; revised manuscript received June 18, 2001; accepted for publication June 28, 2001.

1 Introduction

In recent years there has been a surge of interest in model-based automatic target recognition (ATR) algorithms for use with synthetic aperture radar (SAR) imaging systems. The broad utility of SAR as an imaging methodology is well known, and SAR imaging techniques and systems have been extensively documented.[1,2] The effectiveness of SAR as a remote sensing tool has motivated research into the development of model-based ATR systems.[3,4] Model-based ATR systems identify targets by comparing image features to classification hypotheses generated from a database of physical target models. The generation of target models to populate this database is a problem that is central to the implementation of any model-based ATR system.[3]

We present a framework for producing a three-dimensional target model from the multiple SAR images of a target. Our models consist of spatial collections of reflector primitives such as cylinders, tophats, dihedrals, and trihedrals, each of which is described in terms of a handful of parameters, including a discrete index indicating basic scattering type and several continuous parameters including location, pose, and other information relevant to describing the scattering signature of the overall target.[4,5] Such reflector primitive models offer compact representations of many targets and are well-suited to the feature-driven philosophy of model-based ATR. In particular, reflector primitive models are not only more compact than full facetization or CAD models and potentially more straightforward to generate from SAR imagery, but also allow statistical uncertainty and variability to be modeled directly in feature space. Additionally, such models couple physical relevance to predictive utility in ATR, facilitating the model manipulation and component articulation required to form classification hypotheses.

Our framework entails estimation of the number of scatterers and their descriptive parameters based on the observed set of SAR images. In principle the optimal way to do this is to use all of the available imagery to perform the parameter estimation directly. Note that the explicit inclusion of location as one of the parameters describing each primitive implies that the model estimation procedure must deal with establishing a correspondence between each postulated primitive and the observed scattering responses in all of the SAR images. In principle, the optimal way to do this is to use all of the SAR images directly to establish these correspondences at the same time that the parameters of each primitive are estimated. However, because of the complexity of such a task, the fact that our ultimate objective is a low-dimensional description of the target as a set of primitives, and the fact that model-based ATR systems already operate in this manner, we propose to view the estimation problem as a two-step procedure. Each SAR image is first compressed into a set of augmented detections consisting of relevant information about significant scattering responses in each image, including location and other

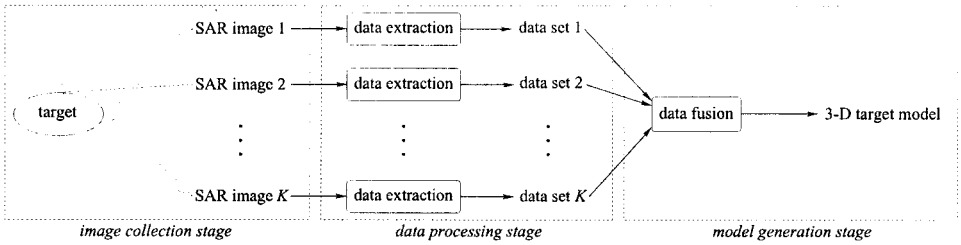


Fig. 1 Target model generation block diagram.

data extracted from the individual images or phase histories. These compressed representations are then fused to estimate the 3-D locations and characteristics of the target primitives. This framework offers great flexibility in the choice of a compression scheme, with possibilities ranging from fine-grained extractions, in which the compression of each SAR image involves keeping a great many basis functions that capture most of the energy in the raw image, to more coarse-grained representations, in which only a small number of dominant scatterers are kept from each image, with only a few parameters describing each response. To introduce our framework and to highlight representations commonly used in ATR, we focus here on a parameterization at the coarser end of this spectrum. This choice also highlights the importance of the correspondence problem mentioned previously.

The major contribution of this paper is to describe a systematic formalism for model generation from multiple SAR images and to provide an initial demonstration of this approach in a constrained environment. In particular, for clarity of presentation, the reflector primitive models we employ here to illustrate and exercise our formalism are of limited extent and variability, and the features we extract from the SAR images are relatively basic. Additionally, there are several places in which we make certain idealizations or opt for a suboptimal implementation for computational reasons. However, the methodology we present not only is extendible to much richer environments, but also provides a clear audit trail for analyzing the effect on algorithm performance and complexity of removing idealizations and approximations. In the next section, we present our formulation of the target model estimation problem, and in Secs. 3 and 4 we describe our application of the expectation-maximization (EM) method to its solution and a modification that enables us to select model order adaptively. In Sec. 5 we present experimental results illustrating the performance of our algorithm. Section 6 concludes with a brief summary and discussion of several of the most significant extensions of our framework to be pursued.

2 Problem Formulation

A block diagram representation of our approach to 3-D target model estimation is depicted in Fig. 1. A target is observed through a set of K SAR images. Each of these images corresponds to a particular viewing geometry, as illustrated in Fig. 2: each image k is characterized by a line-of-sight vector from the center of the synthetic aperture to the center of the target region being imaged. The azimuth ϕ_k and elevation ψ_k defining this line-of-sight vector in terms of a fixed ground frame of reference are arbitrary; we assume each image has been formed at a squint angle of 90°. (Extension of our approach to allow arbitrary squint angles is straightforward.) The synthetic aperture along the platform motion vector and the line-of-sight vector define the slant plane, the imaging plane for the SAR image.[2]

As indicated in Fig. 1, each of the K SAR images is processed to extract a set of observed features, which are then fused to produce a 3-D target model. The framework depicted in Fig. 1 is quite flexible. For instance, we could consider modeling the target as a spatially varying scattering medium and use a trivial data processing stage that simply passes on each complete SAR image to the data processing module, which would then bear the full burden of inverse scattering. However, as indicated in the introduction, we have in mind a much more constrained approach that restricts our description of the target to focus on variables that are most observable and of significant interest for model-based ATR.

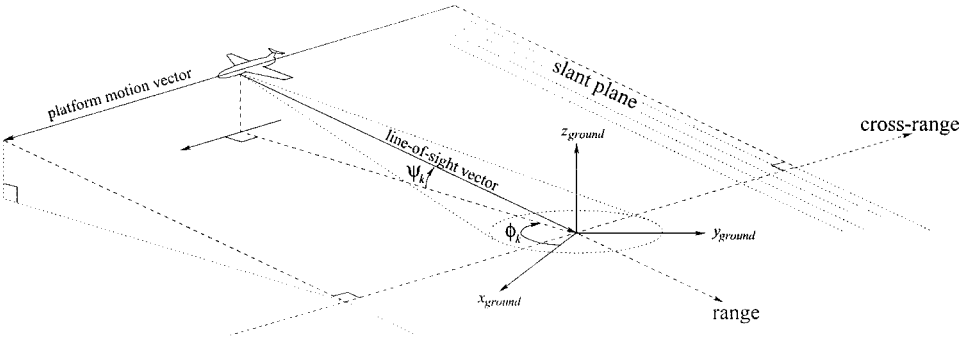


Fig. 2 Imaging geometry and the slant plane for image k.

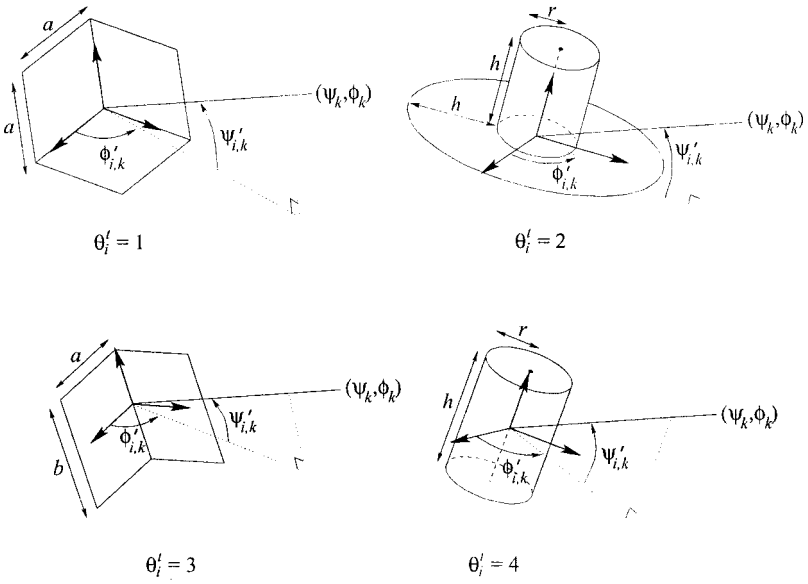


Fig. 3 Reflector primitives: trihedral, tophat, dihedral, and cylinder. Relative elevation $\psi'_{i,k}$ and azimuth $\phi'_{i,k}$ are determined by the absolute viewing elevation ψ_k and azimuth ϕ_k and the pose of primitive, indicated by the orientation of its local axes; primitive dimensions relevant to physical optics RCS models are indicated.

The principal goal is the design of the data fusion module in Fig. 1. This requires specification of exactly what we wish to estimate (i.e., the parameterization of our target models) and how the features serving as input to this module are related to the quantities to be estimated. The latter step involves modeling both the SAR image collection process and the subsequent data processing that produces the observable features on which the fusion module will operate. We describe the notation and basic assumptions defining the problem in Sec. 2.1, and then present our chosen feature extraction procedure in Sec. 2.2. We then present a measurement model in Sec. 2.3 that relates the observable features (the data sets in Fig. 1) and the 3-D target model parameters to be estimated.

2.1 *Target Models: Assumptions and Notation*

Our target models consist of collections of reflector primitives, each of which is described by a small set of parameters that completely specify the scattering behavior of such a primitive given any imaging geometry. As we indicated previously, we restrict attention to a comparatively constrained set of primitives, each of which can be completely described for our purposes by a short vector of parameters. In particular, a target model will be specified by the number of primitives N comprising the target and a vector of parameters $\boldsymbol{\theta}_i$ associated with each component primitive $i = 1,\ldots,N$. In general, we can express this vector as $\boldsymbol{\theta}_i = [\theta_i^t, \theta_i^x, \boldsymbol{\theta}_i^d]$, where θ_i^t is an integer index designating the primitive as one of n_t canonical primitive types, θ_i^x is the 3-D location of the primitive, and $\boldsymbol{\theta}_i^d$ is a generic vector parameter corresponding to a set of continuous-valued descriptors that, along with θ_i^t and θ_i^x, completely specify the log-amplitude scattering response or radar cross section (RCS) of the primitive from any viewing angle.[5,6] (This general parametric formulation for phenomenological modeling of scatterers in SAR images is conceptually similar to those used elsewhere, e.g., in Refs. 7–9. We will denote the log RCS observed from a primitive parameterized by $\boldsymbol{\theta}_i$ and viewed from elevation ψ and azimuth ϕ as $A(\boldsymbol{\theta}_i, \psi, \phi)$, which we typically quote in dBsm. If polarimetric measurements are used, we take the scalar $A(\boldsymbol{\theta}_i, \psi, \phi)$ to be the 2-norm of the magnitude of the polarimetric vector.

We constrain the set of scattering types to a small class of idealized primitives consisting of trihedrals, tophats, dihedrals, and cylinders (so that $n_t = 4$), depicted in Fig. 3; we assign type indices 1 through 4 to these primitives, respectively. For these primitives, $\boldsymbol{\theta}_i^d$ consists of either two or three parameters: an overall base amplitude θ_i^a related to the physical size of the scatterer, a pose θ_i^p indicating the orientation of the scatterer, and a radius of curvature θ_i^r for radially symmetric primitives including tophats and cylinders. Each primitive's location θ_i^x is defined to correspond to the origin of the primitive's local axes, as depicted in Fig. 3. Primitive pose indicates the orientation of these axes with respect to the fixed ground-based coordinate system in terms of three Euler angles.[10] Primitive pose and the absolute viewing angle of image k together define a relative viewing elevation $\psi'_{i,k}$ and azimuth $\phi'_{i,k}$ for each primitive, as depicted in Fig. 3.

The complete vector $\boldsymbol{\theta}_i$ provides a concise yet accurate description of a primitive's appearance in an arbitrary SAR image. Location θ_i^x and radius of curvature θ_i^r (for those primitives for which it is defined), along with the viewing angle, determine the apparent location of the primitive in the slant plane.[1,6] In particular, the apparent location of a primitive in an SAR image is determined essentially by a projection of its apparent 3-D reflection point into the 2-D slant plane. For trihedrals and dihedrals, this apparent reflection point coincides with θ_i^x for most viewing angles;

for radially symmetric primitives such as cylinders and tophats, the apparent reflection point is always uprange from θ_i^x and migrates around the surface of the primitive as the viewing angle varies. We thus model the location of primitive i in image k as

$$\pi_k(\boldsymbol{\theta}_i)=\begin{cases}\mathbf{H}_k\theta_i^x, & \theta_i^t\in\{1,3\},\\ \mathbf{H}_k\theta_i^x-\begin{bmatrix}1\\0\end{bmatrix}\theta_i^r\cos\psi'_{i,k}, & \theta_i^t\in\{2,4\},\end{cases} \tag{1}$$

where $\psi'_{i,k}$ is the relative viewing elevation as pictured in Fig. 3, and where $\mathbf{H}_k$ is the 2×3 ground-to-slant-plane transformation matrix for image k:

$$\mathbf{H}_k=\begin{bmatrix}\cos\psi_k\cos\phi_k & \cos\psi_k\sin\phi_k & -\sin\psi_k\\ -\sin\phi_k & \cos\phi_k & 0\end{bmatrix}. \tag{2}$$

The model of Eq. (1), while accurate for most primitives viewed from most angles, will be inaccurate for primitives viewed at angles at which secondary reflection mechanisms dominate and the apparent specular reflection point does not correspond to θ_i^x (e.g., when a single- or double-bounce response is observed from a trihedral).

The other components of θ_i determine other features of the observed response: discrete type θ_i^t specifies the basic dependence of the response on viewing angle[5] (and, if polarimetric measurements are made, the polarimetric signature vector[11]); pose θ_i^p orients this response by rotating it to correspond to the orientation of the primitive; base amplitude θ_i^a scales the response intensity according to the physical size of the primitive. In particular, physical optics provides expressions for the RCS of each primitive as the product of a size-dependent amplitude term and a unique type-dependent shaping function capturing the dependence of RCS on a relative viewing angle and size.[5] Our log-RCS models are based on these physical optics results and take the form

$$A(\boldsymbol{\theta}_i,\psi_k,\phi_k)=\theta_i^a+S_{\theta_i^t}(\psi'_{i,k},\phi'_{i,k})+c_{\mathrm{pol}}, \tag{3}$$

where θ_i^a encapsulates the fundamental size dependence described by physical optics, where $S_{\theta_i^t}$ is the physical optics log-shaping function describing the variation in scattering response in terms of the viewing angle for all primitives of type θ_i^t, and where c_{pol} is a polarimetry-dependent term that encapsulates any effective gain achieved by using multiple polarimetric channels. For each primitive type, $S_{\theta_i^t}$ is scaled to give a maximum response of 0 dBsm, so that θ_i^a will correspond to the maximum single-polarization RCS of the primitive response. The physical optics models we use for θ_i^a, $S_{\theta_i^t}$, and c_{pol} are detailed in Sec. 7.

One complication in the specification of the $S_{\theta_i^t}$ as in Eq. (3) is the fact that each primitive's physical optics shaping function depends on its dimensions. This dependence is most pronounced for the dihedral and cylinder, which exhibit sinc-like elevation responses depending on b and h, respectively (i.e., those primitives' heights as depicted in Fig. 3). Additionally, primitive responses that comprise multiple reflection mechanisms (e.g., the single-, double-, and triple-bounce response mechanisms of the trihedral) rely on primitive dimensions to determine the relative phase between each component mechanism's response. There are at least three ways in which we can deal with these dependences. We could simply ignore the size dependence and base the $S_{\theta_i^t}$ on nominally chosen dimensions for each primitive, even though real scatterers may have signatures that deviate from these $S_{\theta_i^t}$. Alternatively, we could expand the set of canonical primitives to include several different-sized instances of each basic reflector type. Finally, we could expand the parameterization $\boldsymbol{\theta}_i^d$ to include all relevant dimensions for each primitive type and construct $S_{\theta_i^t}$ with appropriate dependences on these dimensions. For clarity of presentation here we choose the first alternative. In particular, we construct our $S_{\theta_i^t}$ using empirically chosen nominal values for each primitive's size. These nominal values are used to determine the individual responses for each primitive's component reflection mechanisms, which are then combined via a noncoherent sum (i.e., without regard to the size-dependent relative phase) to produce the overall model $S_{\theta_i^t}$ for each primitive type. Extensions to the second or third alternatives listed before are conceptually straightforward (although with a computational cost). In Sec. 5 we explore to what extent a mismatch between actual primitive dimensions and the nominally chosen dimensions selected to construct the $S_{\theta_i^t}$ affects performance.

Before proceeding, we introduce notation $\boldsymbol{\theta}=[\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_N]$. (Note that $\boldsymbol{\theta}$ implicitly specifies the model order N.) Target model generation in our framework is thus estimation of the vector $\boldsymbol{\theta}$ from the data provided by the set of SAR images. We model $\boldsymbol{\theta}$ (and thus N) as unknown parameters about which no information is available other than that provided by the SAR images.

2.2 *Observable Features for Model Generation*

We assume that we have multiple spotlight-mode SAR images[1,2] of the target, formed at arbitrary viewing angles as depicted in Fig. 2. Each of these images is polarimetric, so that a vector measurement $[HH\ \ \sqrt{2}HV\ \ VV]$ is available at each pixel. Furthermore, we assume that all SAR imaging parameters (such as bandwidth, aperture width, range and cross-range locations of each pixel center, and azimuth and depression to the target center) are known, and can be related to the absolute ground-based frame of reference. Such information could be provided, for instance, by geolocation or global positioning measurements taken as the images are collected, coupled with accurate ranging and positioning of the target.

As previously described in the introduction and in conjunction with Fig. 1, we compress the full set of raw SAR imagery by extracting information from each image prior to the model generation stage. For this purpose we utilize a simple amplitude-based peak-extraction technique described in detail in Ref. 12. An arbitrary number of intensity peaks M_k are extracted from each image k; each peak $j=1,\ldots,M_k$ is described in terms of three parameters: a 2-D

slant-plane range/cross-range location $x_{k,j}$, a discrete polarimetric-signature type index $t_{k,j}$, and a scalar log-amplitude $a_{k,j}$. Location and amplitude are obtained using a simple subpixel-interpolation procedure, and polarimetric signature type is obtained via a generalized likelihood ratio test to distinguish between odd-bounce and even-bounce responses.[11] The extracted type is thus a binary variable; we designate an odd-bounce classification as $t_{k,j}=1$, and an even-bounce classification as $t_{k,j}=2$. Note that because trihedrals and cylinders are predominantly odd-bounce scatterers, and dihedrals and tophats predominantly even-bounce scatterers, discrimination between trihedrals and cylinders, or between dihedrals and tophats, is based predominantly on location and amplitude information. The effect of the indistinguishable type measurements for these primitive classes is investigated in Sec. 5.

We note that the performance of our peak extractor depends on the resolution of the SAR imaging system. In particular, not only will the quality of the measurements depend on the resolution of the sensor, but, more fundamentally, our amplitude-based peak-extraction technique cannot distinguish scatterers separated by less than one resolution cell in a given SAR image. Although we could use existing SAR feature-extraction techniques capable of super-resolving scatterers, such as Refs. 8 and 13, we employ the simpler amplitude-based peak-extraction technique to provide a straightforward demonstration of our framework and because to its ease of statistical characterization. Utilization of a different feature extractor would not present any conceptual difficulties.

For convenient reference, the three-parameter location/amplitude/type description of the jth peak of image k is called a report and denoted by $\mathbf{Z}_{k,j}$. At times it will be convenient to refer to the collection of reports within a single image or across images. For these purposes we define notation for all reports in a single image, $\mathbf{Z}_k=[\mathbf{Z}_{k,1},\ldots,\mathbf{Z}_{k,M_k}]$, and notation for all reports in all images, $\mathbf{Z}=[\mathbf{Z}_1,\ldots,\mathbf{Z}_K]$.

2.3 *Measurement Model*

In this section we describe the probabilistic model relating features extracted by the data processor to the target parameters that must be estimated from those features. The uncertainties in the extracted features come at two levels of granularity, one coarse and one fine. The coarse-level uncertainty involves the identity of each measurement: given a set of reports extracted from a single SAR image and a set of target primitives, there is no way of knowing with certainty which reports correspond to which primitives. The fine-level uncertainty involves the stochastic nature of the elements of $\mathbf{Z}_{k,j}$, even given the report's proper correspondence. Compounding the coarse-level uncertainty is the fact that, like any detector, the data processor is subject to missed detections and false alarms, so in general there will not be exhaustive correspondence between the sets of reports and target primitives. To formalize the coarse-level uncertainty we introduce a vector of hidden parameters $\boldsymbol{\lambda}$ that describes the correspondences between reports and target primitives in concrete terms. In particular, we define a label parameter describing the identity of each report $\mathbf{Z}_{k,j}$ as follows:

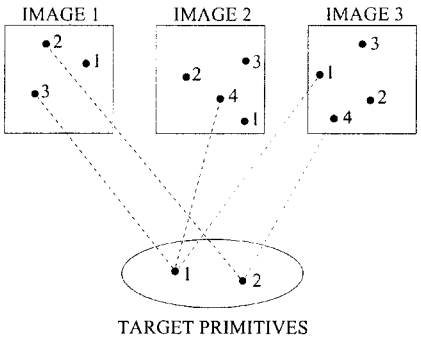


image 1	image 2	image 3
$M_1=3$	$M_2=4$	$M_3=4$
$\lambda_{1,1}=0$ $\lambda_{1,2}=2$ $\lambda_{1,3}=1$	$\lambda_{2,1}=0$ $\lambda_{2,2}=0$ $\lambda_{2,3}=0$ $\lambda_{2,4}=1$	$\lambda_{3,1}=1$ $\lambda_{3,2}=0$ $\lambda_{3,3}=0$ $\lambda_{3,4}=2$
$F_1=1$	$F_2=3$	$F_3=2$

Fig. 4 Notation example.

$$\lambda_{k,j}=\begin{cases} i, & \text{if report } \mathbf{Z}_{k,j} \text{ corresponds to target primitive } i,\\ 0, & \text{if report } \mathbf{Z}_{k,j} \text{ is spurious (corresponds to no primitive).}\end{cases} \tag{4}$$

We also define F_k to be the number of false alarms in image k, i.e., the number of $\lambda_{k,j}$ that equal 0 for a given k. Figure 4 presents an illustrative example of the notation and concepts encapsulated in $\lambda_{k,j}$. This figure depicts a scenario involving two target primitives ($N=2$) and three images ($K=3$).

It is convenient to define a vector $\boldsymbol{\lambda}_k$ collecting the label parameters for all of the reports in image k: $\boldsymbol{\lambda}_k=[\lambda_{k,1},\cdots,\lambda_{k,M_k}]$. The vector $\boldsymbol{\lambda}$ introduced before can be formally defined as $\boldsymbol{\lambda}=[\boldsymbol{\lambda}_1,\cdots,\boldsymbol{\lambda}_K]$. Given knowledge of $\boldsymbol{\lambda}$, the uncertainty remaining in $\mathbf{Z}$ is the distribution of the components of each report $\mathbf{Z}_{k,j}$; this is our fine-level uncertainty. Characterization of the fine-level uncertainty can be done conditionally, and the measurement model can be specified as

$$p(\boldsymbol{\lambda},\mathbf{Z}|\boldsymbol{\theta})=p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta})p(\boldsymbol{\lambda}|\boldsymbol{\theta}), \tag{5}$$

a product of the fine-level probability density function (pdf) and the coarse-level probability mass function (pmf). Throughout this work we describe discrete random variables and vectors such as $\boldsymbol{\lambda}$ by their pmfs, and continuous random variables and vectors such as $\mathbf{Z}$ by their pdfs, using the same notation $p(\bullet)$ in both cases.

We make five general assumptions about the relationship of $\boldsymbol{\lambda}$ and $\mathbf{Z}$ to $\boldsymbol{\theta}$ that facilitate the specification of a measurement model. The first three of these concern the coarse-level uncertainty expressed by $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$; the remaining two concern the remaining fine-scale uncertainty expressed by $p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta})$. These assumptions are largely justifiable on

simple physical grounds, and are standard in a variety of data association contexts.[14,15] These assumptions are as follows.

Assumption 1. False alarms are independent from image to image and do not depend on $\boldsymbol{\theta}$.

Assumption 2. The detectability of the i'th primitive in any image depends only on $\boldsymbol{\theta}_i$ and on the viewing angle of the image; furthermore, missed detections are conditionally independent from image to image and from report to report given $\boldsymbol{\theta}$ and are also independent of false alarms.

Assumption 3. Any primitive generates at most one report in each image, and any report is attributable to at most one primitive.

Assumption 4. Reports in a single image and between images are conditionally independent given $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$, whether they are detections or false alarms.

Assumption 5. The component measurements $\mathbf{x}_{k,j}$, $a_{k,j}$, and $t_{k,j}$ comprising each report are conditionally independent given $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$, whether the report is a detection or a false alarm.

Together, Assumptions 1, 2, and 3 imply the conditional independence of the label parameter vectors for each image:

$$p(\boldsymbol{\lambda}|\boldsymbol{\theta}) = \prod_{k=1}^{K} p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}). \tag{6}$$

Similarly, Assumptions 4 and 5 imply that $p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta})$ can be factored as

$$\begin{aligned} p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta}) &= \prod_{k=1}^{K} p(\mathbf{Z}_k|\boldsymbol{\lambda}_k,\boldsymbol{\theta}) \\ &= \prod_{k=1}^{K} \left[\prod_{j=1}^{M_k} p(\mathbf{x}_{k,j}|\lambda_{k,j},\boldsymbol{\theta})\, p(a_{k,j}|\lambda_{k,j},\boldsymbol{\theta}) \right. \\ &\quad \left. \times p(t_{k,j}|\lambda_{k,j},\boldsymbol{\theta}) \right]. \end{aligned} \tag{7}$$

Although there are situations in which these assumptions will fail—for instance, obstruction will violate Assumption 2, multiple-primitive reflections will violate Assumption 3, and phenomena that cannot be adequately described in the context of the chosen parameterization of Secs. 2.1 and 2.2 could compromise Assumptions 2, 4, and 5—these assumptions are largely realistic and standard.[14,15] In Sec. 6 we discuss how these assumptions might be relaxed to capture more sophisticated real-world effects.

To complete the measurement model, we now need only specify the terms on the right-hand sides of Eqs. (6) and (7). The imagewise identity-uncertainty term $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$ required by Eq. (6) is almost completely determined by Assumption 3 and the constraints it imposes on $\boldsymbol{\lambda}_k$: no more than N of its elements may be nonzero, it cannot contain the same nonzero index twice, and so on. We complete $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$ by assuming a standard Poisson false-alarm model and defining a probability-of-detection function that depends only on a primitive's amplitude in any image k. In particular, we write $P'_{D_{k,i}} \equiv P_D[A(\boldsymbol{\theta}_i,\psi_k,\phi_k)]$, where P_D is a function that we assume is empirically estimated by running the processor on characteristic imagery. If there is no systematic or preferential way of ordering the elements of $\boldsymbol{\lambda}_k$, it then follows that

$$\begin{aligned} p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}) &= \frac{\exp[-\gamma_{FA}V](\gamma_{FA}V)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} (1-P'_{D_{k,i}}) \\ &\quad \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P'_{D_{k,\lambda_{k,j}}}}{1-P'_{D_{k,\lambda_{k,j}}}}, \end{aligned} \tag{8}$$

where V denotes the sensor volume and where γ_{FA} is the false-alarm rate, a parameter that we assume is empirically estimated in the same manner as P_D.

Our models for the report parameter densities of Eq. (7) are conditional on whether the report is a detection or false alarm. False alarms are assumed to be uniformly distributed throughout the SAR images, equally likely to be classified as either polarimetric type, and to have an amplitude distribution denoted by p_{FA}, which can be estimated from characteristic imagery. For a report corresponding to a detection of primitive i, we model $\mathbf{x}_{k,j}$ as a Gaussian with mean $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$ and covariance $\mathbf{R}$, and $a_{k,j}$ as a Gaussian with mean $A(\boldsymbol{\theta}_i,\psi_k,\phi_k)$ and variance σ_a^2, where $\mathbf{R}$ and σ_a^2 can be estimated from characteristic imagery. To model $t_{k,j}$, we assume the availability of an $n_t \times 2$ confusion matrix $\{\rho\}$, where $\rho_{i,j}$ is the probability that the data processor classifies a primitive of type i as having polarimetric signature type j, given that the primitive is detected. As with our other assumed parameters, $\{\rho\}$ can be estimated by processing training data. To simplify notation in subsequent expressions, we write $\rho'_{k,j} \equiv p(t_{k,j}|\lambda_{k,j},\boldsymbol{\theta}) = \rho_{\theta'_{\lambda_{k,j}},t_{k,j}}$ for any detection (i.e., when $\lambda_{k,j}\neq 0$).

We now have all the required components to specify $p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta})$ as in Eq. (7); this can in turn be combined with $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$ according to Eq. (5) to yield a complete measurement model that can be factored into K product terms, one for each image. In particular, $p(\boldsymbol{\lambda},\mathbf{Z}|\boldsymbol{\theta}) = \Pi_{k=1}^{K} p(\boldsymbol{\lambda}_k,\mathbf{Z}_k|\boldsymbol{\theta})$, where

$$\begin{aligned} p(\boldsymbol{\lambda}_k,\mathbf{Z}_k|\boldsymbol{\theta}) &= \frac{\exp[-\gamma_{FA}V]\left(\frac{\gamma_{FA}}{2}\right)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} (1-P'_{D_{k,i}}) \cdot \prod_{j:\lambda_{k,j}=0} p_{FA}(a_{k,j}) \\ &\quad \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P'_{D_{k,\lambda_{k,j}}}}{1-P'_{D_{k,\lambda_{k,j}}}} \cdot \rho'_{k,j} \\ &\quad \cdot \frac{\exp\left\{-\frac{1}{2\sigma_a^2}[a_{k,j}-A(\boldsymbol{\theta}_{\lambda_{k,j}},\psi_k,\phi_k)]^2\right\}}{(2\pi\sigma_a^2)^{1/2}} \\ &\quad \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{\exp\{-\frac{1}{2}[\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}})-\mathbf{x}_{k,j}]^T\mathbf{R}^{-1}[\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}})-\mathbf{x}_{k,j}]\}}{2\pi(\det\mathbf{R})^{1/2}}. \end{aligned} \tag{9}$$

3 A Data Association Approach to Model Generation

The measurement model of Eq. (9) relies on the introduction of a vector of unobservable label parameters $\boldsymbol{\lambda}$ describing the origin of each report. This vector provides not only a convenient device for the specification of a measurement model, but also a conceptual foothold for the estimation of the target parameters. Specifically, if these label parameters were observable—if report data could be associated across images—estimation of $\boldsymbol{\theta}$ would be straightforward. This suggests approaching model generation by way of the underlying data association problem. There is a large body of literature describing theory and methods for solving data association problems in various contexts and applications.[14,16] The chief difficulty facing almost all data association problems, including the one described here, is the combinatorial proliferation of possible correspondences. One way to manage the combinatorial explosion of possibilities is to dismiss as infeasible a majority of associations corresponding to extremely unlikely events; we utilize a technique known as gating, to be described later, for this purpose.[14] Even with such a simplification, however, the remaining data association problem is still formidable and requires a powerful tool for solution. The tool we apply is the expectation-maximization (EM) method.[16,17] In the following section we briefly describe the EM method, and in subsequent sections describe its application to the problem of model generation in the framework we have constructed.

3.1 *Expectation-Maximization Method*

The EM method is an iterative procedure for producing a maximum likelihood (ML) estimate of parameters when there is a many-to-one mapping from a postulated set of ''complete'' data to the set of observed data.[16] In data association problems, the set of complete data comprises the observed data and the vector of associations—$\mathbf{Z}$ and $\boldsymbol{\lambda}$ in our context. Each iteration of the basic EM method consists of two steps: an expectation (E) step and a maximization (M) step. The E step averages the log likelihood of the complete data over all feasible association vectors given the observed data and the latest parameter estimate iterate. The result is an expected log likelihood that is a function of the true parameter vector $\boldsymbol{\theta}$. The M step then maximizes this expected log likelihood with respect to the parameter vector. This yields an estimate of $\boldsymbol{\theta}$ for the current iteration that may be used to recompute the expected log-likelihood in the next iteration's E step. Under relatively mild conditions, the EM method is guaranteed to converge to at least a local maximum of the likelihood function of the observed data.[17,18]

In our context, the EM method proceeds as follows. Let $\boldsymbol{\theta}^{[n-1]}$ be the estimate of $\boldsymbol{\theta}$ produced by the M step in iteration $n-1$. The E step of the n'th iteration requires calculation of the expected log likelihood

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n-1]}) = E[\log p(\boldsymbol{\lambda},\mathbf{Z}|\boldsymbol{\theta})|\mathbf{Z},\boldsymbol{\theta}^{[n-1]}] = \sum_{\boldsymbol{\lambda}\in\Lambda} [\log p(\boldsymbol{\lambda},\mathbf{Z}|\boldsymbol{\theta})]p(\boldsymbol{\lambda}|\mathbf{Z},\boldsymbol{\theta}^{[n-1]}), \quad (10)$$

where Λ is the set of all possible $\boldsymbol{\lambda}$. The M step then requires maximization of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n-1]})$ over $\boldsymbol{\theta}$. Specifically, the M step determines the n'th iterate value:

$$\boldsymbol{\theta}^{[n]} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n-1]}). \quad (11)$$

We describe the implementation of the E and M steps for our problem in Secs. 3.2 and 3.3, respectively. Because the EM method is an iterative procedure, it requires an initialization $\boldsymbol{\theta}^{[0]}$ and a criterion for termination; the latter is discussed in Sec. 3.4, while our initialization procedure is described in Sec. 4 In addition, while in principle the EM method can deal directly with unknown model order N by introducing this quantity as part of the complete data, this adds an undesirable level of complexity to our algorithm. As a result, the basic algorithm we develop in Secs. 3.2 and 3.3 assumes that N is specified. In Sec. 4 we describe a modification to the basic algorithm that enables adaptive selection of model order as the algorithm progresses.

3.2 *Implementation of the E Step*

It follows from Eq. (9) that with N specified, the expected log likelihood to be calculated in the E step as in Eq. (10) can be expressed as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_{i=1}^{N} Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) + C_K = \sum_{i=1}^{N}\sum_{k=1}^{K} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) + C_K, \quad (12)$$

where

$$Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{k=1}^{K} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) \quad (13)$$

and

$$\begin{aligned} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{j=1}^{M_k} \Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) \\ \times\Bigg[\log\frac{P'_{Dk,i}}{1-P'_{Dk,i}} + \log\rho'_{k,j} \\ -\frac{1}{2\sigma_a^2}[a_{k,j}-A(\boldsymbol{\theta}_i,\psi_k,\phi_k)]^2 \\ -\frac{1}{2}[\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)]^T\mathbf{R}^{-1}[\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)]\Bigg] \\ +\log(1-P'_{Dk,i}). \end{aligned} \quad (14)$$

In other words, the expected log likelihood separates into NK terms, each of which depends only on a single target primitive, and the M_k reports extracted from the given image k. This decoupling of the expected log likelihood is a consequence of our independence assumptions of Sec. 2.3. (A similar decomposition will be possible in the M step.)

This is an encouraging result, because it means that the computational complexity of the E step will increase only linearly with K and N.

Examining Eq. (14), we see that the computation of the E step uses quantities specified previously, as well as the report-to-primitive association probabilities $\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$. In theory these probabilities can be calculated via Bayes' rule. In practice, however, this computation is typically intractable even for problems of modest size due to the combinatorial dependence of the number of possible $\boldsymbol{\lambda}_k$ on N and M_k. (For the multiple-primitive example of Sec. 5, the set of possible association vectors typically numbers in the hundreds; for a problem involving as few as a dozen primitives, the cardinality increases to billions.) To overcome this difficulty, we use a common and easily justifiable simplification known as gating.[14] Specifically, complete enumeration of the set of possible $\boldsymbol{\lambda}_k$ entails consideration of all possible associations, even very unlikely ones in which measured locations $\mathbf{x}_{k,j}$ are associated with target primitives that project to points in the slant plane far from $\mathbf{x}_{k,j}$. Gating is a method for excluding such unlikely pairings from consideration by adaptively defining the set of feasible associations to be the much smaller set of $\boldsymbol{\lambda}_k$ that correspond to associations between reports and primitives that are believed to be close enough, i.e., for which $\|\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}})\|_2 \leq r_{\text{gate}}$, $j=1,\ldots,M_k$. Typically r_{gate} is taken as a small multiple of $[\text{trace}(\mathbf{R})]^{1/2}$.

3.3 *Implementation of the M step*

The M step requires maximization of the E step's expected log likelihood $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ with respect to $\boldsymbol{\theta}$ as in Eq. (11). The separation of this expected log likelihood into independent terms for each primitive in Eq. (12) implies that this maximization may be achieved independently for each primitive. In particular, the M step requires N independent maximizations, each of a single $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\boldsymbol{\theta}_i$. Since $\boldsymbol{\theta}_i$ includes both continuous parameters (θ_i^{x}, θ_i^{p}, θ_i^{a}, and possibly θ_i^{r}) and a discrete parameter (θ_i^{t}), we are faced with a hybrid maximization problem for each primitive, with the discrete parameter limited to a small, finite space of n_t elements. We thus maximize $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ by performing n_t separate trial maximizations over the continuous components of $\boldsymbol{\theta}_i$, one for each possible value of θ_i^t. Examination of Eq. (14) reveals that each trial maximization is nontrivial: there is a complicated relationship between $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ and the set of continuous parameters. Specifically, the pose, location, and radius terms are coupled due to $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$, and the pose and base amplitude are coupled due to $P_{Dk,i}'$ and $A(\boldsymbol{\theta}_i,\psi_k,\phi_k)$.

Consider the following approximate maximization over the continuous parameters with θ_i^t fixed, equivalent to a single-iteration coordinate ascent. First, maximize $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over pose while fixing amplitude, location, and radius at their maximizing values from the previous iteration; this can be accomplished with a coarse-to-fine search over pose. Second, perform a line search to maximize over amplitude with pose fixed at the value just obtained, and with location and radius fixed at their values from the previous iteration. Finally, maximize $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over location and radius while fixing pose and amplitude at the values just obtained; this can be done in closed form as a weighted least-squares error (WLSE) estimate based on report locations (where the E-step report-to-primitive correspondence probabilities give the weights) due to the quadratic dependence of $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ on location and radius. In particular, for $\theta_i^{t[n]}\in\{1,3\}$ we have

$$\theta_i^{\mathbf{x}[n+1]}=\left[\sum_{k=1}^{K}\sum_{j=1}^{M_k}\mathbf{H}_k^T\mathbf{R}^{-1}\mathbf{H}_k\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\right]^{-1}\times\left[\sum_{k=1}^{K}\sum_{j=1}^{M_k}\mathbf{H}_k^T\mathbf{R}^{-1}\mathbf{x}_{k,j}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\right], \tag{15}$$

and we can obtain a similar expression to maximize over θ_i^{x} and θ_i^{r} when $\theta_i^{t[n]}\in\{2,4\}$.[12] [If the WLSE estimate gives a negative radius in this case, we set the radius estimate to zero and use Eq. (15) for the location estimate.] This type of partitioned M-step implementation is known as expectation conditional maximization (ECM) and is sufficient to ensure eventual convergence of the EM method to a maximum of the likelihood function under the same conditions as an algorithm that achieves a true joint maximum at each M step.[16] If not for the pose search, the computational burden of the M step would generally be insignificant compared to that of the E step. As it is, however, the M step greatly exceeds the E step in execution time.

3.4 *Termination Criterion*

Rather than directly monitoring $p(\mathbf{Z}|\boldsymbol{\theta}^{[n]})$ for convergence, we adopt the computationally simpler and widely used procedure of monitoring the estimates $\boldsymbol{\theta}^{[n]}$ themselves. Once the estimates of θ_i^t produced by the M step remain fixed between iterations and the changes in the continuous parameter estimates all drop below specified thresholds, the iteration is terminated and the final $\boldsymbol{\theta}^{[n]}$ is used as the final estimate of $\boldsymbol{\theta}$.

4 Initialization and Model Order Estimation

In this section we describe two final features of our model generation algorithm, namely the initialization stage and a modification to the standard EM iteration that enables adaptive model order selection as the iteration progresses. Our model order adjustment stage occurs after every M step, before the termination criterion check, and is capable only of reducing the model order or leaving it unchanged. This imposes the important guideline that the initialization should be biased toward overestimating N: any overfit can be corrected in subsequent iterations by the model order reduction stage, but any underfit is permanent. We describe our initialization procedure in Sec. 4.1, and model order reduction in Sec. 4.2.

4.1 *Initialization of the EM Method*

Our initialization procedure is based on an agglomerative clustering algorithm that groups reports between images based on the set of all $\mathbf{x}_{k,j}$. Each group of reports produced by this agglomeration is used to initialize a single target

primitive. Because the number of clusters produced by the agglomeration depends on the data, the order of $\boldsymbol{\theta}^{[0]}$ is adaptive to $\mathbf{Z}$. As we indicate later, we can bias toward overfitting by setting an agglomeration threshold appropriately. Our initialization procedure is completely described in Ref. 12; the following is an outline of the procedure that omits some minor details and modifications made for computational reasons.

Agglomerative clustering is a general procedure in which a collection of items is iteratively grouped into distinct clusters by successively merging items and groups based on some measure of their dissimilarity.[19] At each iteration, the two most similar items or groups in the collection are merged. This is repeated until all remaining groups are more dissimilar than some threshold η. The number of clusters produced depends strongly on η: a large choice for η will generally result in fewer clusters than a small choice.

Our dissimilarity measure is a chi-squared statistic that is based on the measured report locations $\mathbf{x}_{k,j}$ and motivated by the following observation: reports from different images that could be explained by projection from a single point or closely spaced points in $\mathbb{R}^3$ are likely to correspond to the same target primitive. Reports that cannot be explained by closely spaced points in $\mathbb{R}^3$ (or reports from within the same image) are dissimilar and should not be grouped. It can be shown[12] that if a group of reports all correspond to the same primitive according to the Gaussian model described in Sec. 2.3 with an identity-multiple covariance $\mathbf{R}$, then the total squared error of the linear least-squares error (LLSE) estimate of the corresponding primitive's 3-D location (i.e., the sum of the squared 2-D distances between the projections of the LLSE estimate and the report locations) is a chi-squared random variable. If the reports come from multiple primitives, the total squared error will in general be much larger and will have a noncentral chi-squared distribution. If the reports include false alarms, which we model as having uniform rather than Gaussian distributions, the total squared error will have neither a central nor a noncentral chi-squared distribution, but will still generally be larger than when all reports are attributable to a single primitive. We thus use the chi-squared cumulative distribution function as our dissimilarity measure. This provides an interpretation of η as a confidence level for accepting or rejecting a candidate group based on how well a single primitive could explain the report locations. We set η to a value empirically determined to give a low probability of underfitting, so that the initialization is likely to contain more hypothesized primitives than are necessary to describe the target. Some of these primitives will be removed in the subsequent model order reduction stage.

We use each report cluster produced by the agglomeration to initialize a single target primitive parameter vector $\boldsymbol{\theta}_i$ as follows. Base amplitude is set to the maximum log amplitude among the clustered reports, minus c_{pol} of Eq. (3). Type and pose are chosen to maximize the likelihood of having observed the amplitudes and types in the report group, given that all the reports correspond to a single primitive and given the just-selected base amplitude. (Type and pose maximization are achieved by a search similar to that described in Sec. 3.3 in conjunction with the M step.) Location (and radius, if deemed applicable by the type initialization) is estimated by a least-squares approximation based on the set of $\mathbf{x}_{k,j}$ similar to that of Eq. (15).

The agglomeration described here is based solely on location, although it can easily be extended to include type and amplitude information. In addition, note that the procedure we have described is not perfectly designed for grouping reports produced by cylinders and tophats since, as is evident from Eq. (1), the apparent reflection point for such primitives changes from view to view due to the radial symmetry of these reflectors. As a result, the dissimilarity values computed by our clustering method for a group of reports produced by such a primitive will generally be larger than those computed for trihedrals or dihedrals. While it is possible to design a more sophisticated clustering method to deal with this effect, we have found our simpler algorithm to be sufficient to produce a satisfactory initialization for the EM iteration.

4.2 *Model Order Reduction*

Our model order reduction stage counteracts the overfitting induced in the initialization. This is accomplished by examining the empirical evidence for each hypothesized primitive's existence after each M step and removing any primitives whose estimates have converged to values which are not strongly supported by the data. More precisely, after each M step we calculate an empirical probability of detection $\tilde{P}_{D:i}$ for each primitive i for which $\boldsymbol{\theta}_i^{[n]}$ has converged:

$$\tilde{P}_{D:i}=\frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{M_k}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}). \tag{16}$$

[The $\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ terms are available from the E step of the iteration.] Intuitively, if $\tilde{P}_{D:i}$ is near zero, then $\mathbf{Z}$ provides little evidence to support the hypothesis of a target primitive whose parameters are given by $\boldsymbol{\theta}_i^{[n]}$. This suggests that $\boldsymbol{\theta}_i^{[n]}$ is either a poor estimate (possibly due to convergence of the EM iteration to a local maximum of the likelihood function) or that the model order is too high. In either case it is appropriate simply to remove the primitive from the estimate and decrement the model order. We make this decision by comparing the $\tilde{P}_{D:i}$ of the converged primitives to a type-dependent threshold (from a set of n_t thresholds empirically chosen to reflect the fact that primitives with highly specular responses will almost certainly be detected in fewer images than those with spatially persistent responses). This model order reduction stage ensures that all primitives contained in the final estimate of θ will have empirical probabilities of detection at least as great as their type-dependent thresholds dictate. One drawback to our approach is the increased computational burden in the early stages of the EM iteration, before the extraneous primitives have been eliminated and N is still artificially large. The benefit of this extra computation is the increased adaptivity and robustness of the final estimate θ.

5 Results

In this section, we present results of the application of our algorithm to synthetic SAR imagery generated by XPatch,

Table 1 Measurement model parameters.

quantity	notation	value
location covariance	$\mathbf{R}$	$(5.0\text{cm})^2 \times \mathbf{I}$
normalized false alarm rate	γ_{FA}	$0.058 / \text{m}^2$
type confusion matrix	$\{\rho\}$	$\begin{bmatrix} 0.78 & 0.01 & 0.17 & 0.97 \\ 0.22 & 0.99 & 0.83 & 0.03 \end{bmatrix}^T$
amplitude variance	σ_a^2	$(5\text{dBsm})^2$

Table 2 Single-primitive targets: primitive dimensions.

primitive	θ_i^a [dBsm]	dimensions
trihedral	10	$a = 12.67$ cm
tophat	10	$r = 18.39$ cm, $h = 36.78$ cm
dihedral	10	$a = 14.05$ cm, $b = 14.05$ cm
cylinder	10	$r = 17.70$ cm, $h = 53.05$ cm

an electromagnetic simulation package capable of accurately simulating arbitrary electromagnetic scattering measurements obtained by interrogating a facetization-model target with radiation.[20,21] We use the XPatch-T module of the package to produce image chips at a range and cross-range resolution of 0.3 m, a range and cross-range pixel spacing of 0.2 m, and a center frequency of 9.6 GHz, and use a -35-dB Kaiser sidelobe weighting function for image formation from phase history data. XPatch produces an image of a target in the absence of natural clutter; we model clutter as an additive *K*-distributed process independent for each pixel, with grassy terrain parameters as reported in Ref. 22.

Recall that the measurement model described in Sec. 2.3 is parameterized by several quantities that must be specified in advance. The quantities we use for the experiments in this section are given in Table 1. The location covariance, false alarm rate, and confusion matrix given in Table 1 are average values compiled by processing a set of training images, each containing a single primitive in the grassy terrain clutter environment. The probability-of-detection function and false-alarm amplitude pdf are histograms compiled from the training results. The amplitude variance term is a heuristic value chosen with the intention of capturing some of the variability in primitive responses encountered in the real world that would be difficult to model in a training set (e.g., geometrical deviations or perturbations from ideality). Recall also that construction of the primitive scattering models requires specification of relevant primitive dimensions, as described in Secs. 2.1 and 7. For most of the results in this section (i.e., all results in Secs. 5.1, 5.2, and 5.3), the scattering-model primitive dimensions have been set to the dimensions of the primitives to be estimated; in Sec. 5.4 we examine the effects of primitive dimension mismatch on algorithm performance.

For each target described, we generated a superset of 2736 XPatch images—one for each viewing angle on a 2.5-deg elevation/azimuth grid extending in elevation from 5 to 50 deg and in azimuth from 0 to 357.5 deg. For each Monte Carlo run, we selected a random subset of images to give an average view sampling density (AVSD) of either 10 or 20 deg, i.e., a random subset containing exactly enough images to create a 10-deg grid (180 images) or a 20-deg grid (54 images) if the images were equally spaced, and corrupted each image with independent *K*-distributed clutter as described earlier.

5.1 *Single-Primitive Targets*

Our first set of experiments details the performance of the algorithm on four targets, each consisting of a single primitive with a base amplitude of 10 dBsm. This serves to demonstrate the relative observability of the four primitive types and to establish a rough benchmark for results from more complex targets. Each of the four targets in this section corresponds to a single primitive (a unique type for each target) located at ground coordinates [30.5 0 15.2 cm]. The trihedral and tophat are oriented with their bases parallel to the ground plane, the trihedral rotated to give a maximum specular response at azimuth 0 deg; the dihedral and cylinder are oriented so that a maximum specular response is obtained at elevation 25 deg and azimuth 0 deg. Primitive dimensions (see Fig. 3) are given in Table 2. (Choosing primitive sizes to give a base amplitude of 10 dBsm facilitates comparison of primitive observability but requires the dihedral and trihedral to be much smaller than typically observed dihedrals and trihedrals in many targets of interest.[6])

Tables 3 and 4 present the performance of the algorithm on these four targets for both the 10- and 20-deg AVSDs. Table 3 presents model order and type estimation statistics: P_{det} refers to the fraction of runs in which an estimate was produced for the primitive, i.e., in which it was captured by the initialization stage and survived the model order reduction stage through convergence of the EM iteration; overfit refers to the fraction of runs in which the final model order was greater than one. Type confusion presents type estimation results conditional on detection. In most runs in which the primitive is detected and estimated to be the correct type, the continuous parameter estimates cluster near the true values; in a handful of runs they do not. The fraction of runs in which this occurs is listed in the fraction spurious column. Table 4 (described later) presents the results of the continuous parameter estimation, conditional on correct type identification and nonspurious parameters. For each of the eight experiments (four primitives, two AVSDs), Monte Carlo runs were continued until we had obtained 500 trials in which the primitive was detected and estimated to be the correct type.

Table 3 illustrates the relative observability of the primitive types. Trihedrals and tophats have broad angular responses and are easily detected by the algorithm; dihedrals and cylinders have responses largely confined to a single azimuthal plane and are more difficult to detect. The dihedral is much shorter than the cylinder (see Table 2) and thus has a much broader out-of-plane elevation response,[5,6] accounting in part for the greater detectability of the dihedral. It is also apparent from Table 3 that the type classification performance of the algorithm is excellent: in almost every trial in which the primitive is detected, its type is correctly identified. This suggests that the limited type information provided by the even-bounce/odd-bounce discriminator in the data extraction stage, as discussed in Sec. 2.2, is not a significant impediment to type estimation.

Table 3 Single-primitive targets: model order and type confusion statistics.

AVSD	target	model order		type confusion				fraction spurious
		P_{det}	overfit					
10 deg	trihedral	1.000	0.000	[1.000	0.000	0.000	0.000]	0.000
10 deg	tophat	1.000	0.000	[0.000	1.000	0.000	0.000]	0.000
10 deg	dihedral	0.721	0.000	[0.000	0.000	0.998	0.002]	0.002
10 deg	cylinder	0.490	0.000	[0.000	0.000	0.021	0.979]	0.002
20 deg	trihedral	0.948	0.000	[0.990	0.000	0.006	0.004]	0.000
20 deg	tophat	1.000	0.000	[0.000	1.000	0.000	0.000]	0.000
20 deg	dihedral	0.185	0.000	[0.035	0.000	0.933	0.032]	0.000
20 deg	cylinder	0.109	0.002	[0.030	0.000	0.004	0.966]	0.026

Table 4 presents the error statistics for the continuous parameters. Base amplitude figures are quoted in dBsm, location and radius in inches, and pose in degrees. We break up location and base amplitude statistics into bias and covariance to illustrate bias effects described later. We separate pose results into two components because, although in general three Euler angles (corresponding essentially to elevation, azimuth, and rotation[10]) are required to specify the pose of a primitive, only two angles (azimuth and elevation) are required to specify the pose of the tophat and cylinder due to those primitives' rotational symmetry. The azimuth/elevation root-mean-squared error (RMSE) statistics correspond to angular separation in degrees between two points on a sphere.

The amplitude bias exhibited in Table 4 is largely attributable to two factors. First is the frequency windowing inherent in the SAR imaging process.[1,2] In particular, a primitive's brightness in an image is affected by its location in the slant plane relative to the pixel centers. In most images a detected primitive will not project directly onto a pixel center, resulting in a tendency to underestimate amplitude. In the absence of other effects, it can be shown[12] that this nominal bias is roughly -0.5 dBsm. (It is possible to correct for this factor simply by adding 0.5 to the amplitude estimates produced by the algorithm, although we have not done this in the presentation of the results in Table 4.) The second factor influencing amplitude bias is the slight mismatch between $S_{\theta_i^t}$ and the XPatch-predicted scattering responses, which is attributable to several sources including the formation of the $S_{\theta_i^t}$ as noncoherent sums of individual response mechanisms as described in Secs. 2.1 and 7, inaccuracies in the physical optics approximations for basic scattering mechanisms,[5,6] and XPatch-related shooting-and-bouncing-rays and facetization effects.[20,21] (A more detailed analysis of these factors is presented in Ref. 12.) The base amplitude standard deviation figures indicate that, as expected, the greater observability afforded by a broad angular response results in a lower standard deviation. The same trend holds for the location and radius statistics.

The trihedral and dihedral location estimates exhibit a pronounced bias. This is due to the influence of these primitives' lower-bounce mechanisms described earlier. For instance, although trihedral triple-bounce reflections all appear to emanate from that primitive's apex at θ_i^x (see Fig. 3), trihedral double-bounce reflections appear to emanate from the trihedral creases, and single-bounce reflections appear to emanate from the trihedral plates. The WLSE location estimate of Eq. (15) [induced by the model of Eq. (1)] is biased by these lower-bounce reflections. In Sec. 5.2 we discuss an approach to removing this bias, if desired. Additionally, although it is not illustrated in Table 4, the radius errors are correlated with the location errors due to layover effects.

Table 4 Single-primitive targets: base amplitude, location, pose, and radius statistics.

AVSD	target	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^x$ error [cm]		$\hat{\theta}_i^p$ rmse [°]		$\hat{\theta}_i^r$ rmse [cm]
		bias	stdev	$\|\text{bias}\|$	$\sqrt{\text{tr(cov)}}$	az/el	rot	
10 deg	trihedral	0.156	0.494	6.233	3.025	2.169	7.786	—
10 deg	tophat	−1.382	0.191	0.409	1.288	1.329	—	1.056
10 deg	dihedral	−0.148	1.114	11.511	14.171	10.592	7.471	—
10 deg	cylinder	0.080	1.625	0.410	22.169	1.286	—	6.629
20 deg	trihedral	−0.000	1.035	5.559	6.831	9.175	22.363	—
20 deg	tophat	−1.298	0.365	0.446	2.175	2.388	—	1.730
20 deg	dihedral	−0.052	1.840	7.591	31.417	14.285	11.644	—
20 deg	cylinder	0.455	2.771	1.786	30.845	3.433	—	9.980

Table 5 Single-primitive targets: trihedral and dihedral location estimate refinement.

		unrefined location error [cm]		refined location error [cm]	
AVSD	primitive	$\|\text{bias}\|$	$\sqrt{\text{tr}(\text{cov})}$	$\|\text{bias}\|$	$\sqrt{\text{tr}(\text{cov})}$
10 deg	trihedral	6.233	3.025	0.576	4.792
10 deg	dihedral	11.511	14.171	3.259	18.012
20 deg	trihedral	5.559	6.831	0.735	12.652
20 deg	dihedral	7.591	31.417	3.671	38.742

Table 6 Multiple-primitive target: primitive locations and dimensions.

primitive	θ_i^{xT} [cm]			θ_i^a [dBsm]	dimensions
trihedral	[−45.7	−45.7	0]	23.75	$a=28.00$ cm
tophat	[−45.7	45.7	0]	10	$r=18.39$ cm, $h=36.78$ cm
dihedral	[45.7	45.7	0]	25	$a=33.30$ cm, $b=33.30$ cm
cylinder	[45.7	−45.7	0]	10	$r=17.70$ cm, $h=53.04$ cm

The pose results of Table 4 demonstrate that primitive pose can in general be accurately estimated to a finer granularity than provided by the AVSD. The dihedral pose errors, which are larger than those observed for the other primitive types, are attributable to the near-invariance of the dihedral response to certain changes in the Euler-angle pose.[12] The relatively large trihedral rotational error stems from the near-invariance of the trihedral response to changes in the rotation Euler angle.[5,6]

5.2 *Reducing Trihedral and Dihedral Location Estimate Bias*

As described in the previous section, the location estimates of the trihedral and dihedral, formed according to Eq. (15) at each iteration, are biased. The direction and magnitude of the bias will depend on the primitive orientation and dimensions. If we seek an unbiased location estimate, we could modify the model of Eq. (1) to take into account the different reflection points of the lower-bounce mechanisms. Unfortunately, this introduces a dependence on primitive dimension and pose, and complicates the maximization at each iteration. We opt for a computationally simpler approach that is in effect a postprocessing step to be implemented after the EM iteration has converged. Specifically, at the conclusion of the iteration we have available a final estimate of primitive pose (from the final M step) and a final set of report-to-primitive correspondence probabilities (from the final E step). We use the E-step probabilities and M-step pose estimate to form a new WLSE estimate of primitive location based only on those reports from relative viewing angles at which the highest bounce response dominates, as deemed by the M-step pose estimate. If this pose estimate is accurate, this will correct the bias (though potentially at the cost of a higher overall RMS location error).

The results of this postprocessing on the trials of the previous section are presented in Table 5. The trihedral location estimation refinement achieves a greater fractional reduction in bias than the dihedral location estimate refinement due to the greater accuracy of the trihedral pose estimates (see Table 4). For similar reasons, the refinement is more successful for the 10-deg AVSD than for the 20-deg AVSD.

5.3 *Multiple-Primitive Target*

The results of Sec. 5.1 demonstrate the performance of the algorithm on single-primitive targets, for which the underlying data association problem faced by the algorithm is quite simple. In this section, we examine algorithm performance when faced with a more challenging data association problem, specifically for a multiple-primitive target. We would expect the error statistics obtained for each primitive in a multiprimitive target to be no better than those of Sec. 5.1, which were obtained in the absence of confusing primitives. Similar statistics would indicate that the algorithm is successfully solving the implicit data association problem.

The multiple-primitive target we examine here consists of four primitives, one of each type, centered at the corners of a 91.4-cm square. The dimensions and locations of these primitives are given in Table 6. The cylinder and tophat are the same size as those of Sec. 5.1; the dihedral and tophat are larger (but still smaller than many dihedrals and trihedrals that might be encountered in targets of interest). Be-

Table 7 Multiple-primitive target: model order and type confusion statistics.

AVSD	target component	P_{det}	type confusion				overfit	fraction spurious
10 deg	trihedral	1.000	[1.000	0.000	0.000	0.000]	0.002	0.008
10 deg	tophat	1.000	[0.000	1.000	0.000	0.000]		
10 deg	dihedral	0.950	[0.000	0.000	0.998	0.002]		
10 deg	cylinder	0.406	[0.000	0.000	0.045	0.955]		
20 deg	trihedral	1.000	[1.000	0.000	0.000	0.000]	0.000	0.016
20 deg	tophat	1.000	[0.000	1.000	0.000	0.000]		
20 deg	dihedral	0.460	[0.026	0.004	0.953	0.017]		
20 deg	cylinder	0.082	[0.024	0.000	0.024	0.952]		

Table 8 Multiple-primitive target: base amplitude, location, pose, and radius statistics.

AVSD	target component	$\hat{\theta}_i^a$ error [dBsm] bias	$\hat{\theta}_i^a$ error [dBsm] stdev	$\hat{\theta}_i^x$ error [cm] $\|\text{bias}\|$	$\hat{\theta}_i^x$ error [cm] $\sqrt{\text{tr(cov)}}$	$\hat{\theta}_i^p$ rmse [°] az/el	$\hat{\theta}_i^p$ rmse [°] rot	$\hat{\theta}_i^r$ rmse [cm]
10 deg	trihedral	−0.617	0.495	7.635	2.395	2.978	17.866	—
10 deg	tophat	−1.119	0.222	0.339	1.407	1.810	—	1.126
10 deg	dihedral	−1.696	2.411	7.227	12.647	11.337	6.934	—
10 deg	cylinder	0.154	1.507	1.190	21.713	1.209	—	7.553
20 deg	trihedral	−0.900	0.959	7.236	4.821	4.911	25.839	—
20 deg	tophat	−1.026	0.374	0.531	2.480	2.683	—	1.923
20 deg	dihedral	−2.873	4.045	7.368	21.776	13.483	15.391	—
20 deg	cylinder	0.517	2.567	2.718	29.500	1.869	—	7.480

cause the cylinder and tophat are the same size as those of Sec. 5.1, we can directly compare the results for these primitives between the two sections.

We performed 500 Monte Carlo runs for both a 10-deg AVSD and a 20-deg AVSD for the multiple-primitive target. The results are presented in Tables 7 and 8 in a format similar to that of Tables 3 and 4. (Overfit and spurious-estimate results in Table 7 are presented for the target in ensemble rather than for individual primitives.) The tophat and cylinder results here compare favorably to those of Sec. 5.1, suggesting that the algorithm is overcoming the more challenging data association problem. In particular, tophat detection and type-estimation performance are unaffected, and tophat continuous-parameter error statistics are only slightly worse; cylinder statistics are generally of the same quality or slightly worse than those observed in Sec. 5.1. (For the 20-deg AVSD experiment in particular, the small number of detections limits the statistical significance of the cylinder results.) Most marked is the decrease in cylinder detectability in the 10-deg AVSD experiment, suggesting that the initialization suffers from the presence of more primitives.

Comparison of the dihedral and trihedral statistics here to those of Sec. 5.1 demonstrates the impact of greater primitive observability. One marked change is the improved detectability of the larger dihedral. Also apparent are the smaller location covariances of both primitives. (The location estimates can be refined according to the procedure described in Sec. 5.2; the results of that refinement are displayed in Table 9.) Pose errors have slightly worsened for the dihedral because dihedral size does not appreciably affect the dihedral response's pose invariance (described in Sec. 5.2). The trihedral rotational errors have worsened because the larger trihedral dimensions here make the single- and double-bounce reflections less prominent and the overall response more invariant to changes in rotation. Finally, the larger amplitude biases and standard deviations seen here are attributable to the initialization of base amplitude as a maximum report amplitude, which tends to underestimate base amplitude. In some runs this underestimate persists to convergence.

Table 9 Multiple-primitive target: trihedral and dihedral location estimate refinement.

AVSD	primitive	unrefined location error [cm] $\|\text{bias}\|$	unrefined location error [cm] $\sqrt{\text{tr(cov)}}$	refined location error [cm] $\|\text{bias}\|$	refined location error [cm] $\sqrt{\text{tr(cov)}}$
10 deg	trihedral	7.635	2.395	2.448	2.288
10 deg	dihedral	7.227	12.647	4.729	11.750
20 deg	trihedral	7.236	4.821	2.895	4.487
20 deg	dihedral	7.368	21.776	4.358	24.013

5.4 *Primitive Dimension Mismatch*

The scattering models used to generate the results of the previous sections were all constructed using primitive dimensions perfectly matched to the primitives being estimated. In practice this is not possible. A model generation algorithm will have to contend with unknown dimensions and multiple-sized instances of each primitive class. As described in Sec. 2.1, there are several ways to deal with this issue, ranging from the method we have chosen to the conceptually optimal but computationally demanding approach of including all relevant primitive dimensions in the target parameterization. In this section, we examine the impact of primitive dimension mismatch on algorithm performance and address the question of whether the nominal-dimension approach taken here is sufficient to envelop the range of phenomenology observed from primitives of different sizes.

All four primitive scattering functions depend on primitive dimensions (see Sec. 7). This dependence is most marked for the dihedral and cylinder, whose primary response mechanisms are fundamentally affected by primitive height; the dependence is relatively minor for trihedrals

Table 10 Dimensions of three dihedrals.

dihedral	θ_i^a [dBsm]	dimensions
short	25	$a=53.04$ cm, $b=20.90$ cm
square	25	$a=33.30$ cm, $b=33.30$ cm
tall	25	$a=20.90$ cm, $b=53.04$ cm

Table 11 Matched primitive dimensions: dihedral estimation results.

			$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^x$ error [cm]		$\hat{\theta}_i^p$ rmse [°]	
dihedral	P_{det}	P_{id}	bias	stdev	$\|\text{bias}\|$	$\sqrt{\text{tr(cov)}}$	az/el	rot
short	1.000	1.000	−3.218	1.804	7.394	12.781	10.394	8.274
square	0.953	0.996	−2.033	2.338	5.764	13.890	11.713	9.281
tall	0.822	0.960	−1.242	2.746	10.593	11.896	13.675	7.714

and tophats, whose primary response mechanisms do not vary with primitive dimension. We limit our investigation here to the effects of dihedral dimension mismatch. The effects of cylinder dimension mismatch can be expected to be similar, due to the similar dependence of the dihedral and cylinder shaping functions on primitive dimension. The effects of mismatch on trihedral and tophat estimation performance can be expected to be less marked than those observed for the dihedral; due to the trihedral's and tophat's secondary dependence (i.e., only observed in the lower bounce mechanism responses) on primitive size.

Table 10 presents dimensions of three 25-dBsm dihedrals (short, square, and tall), each with different plate dimensions. To establish a benchmark against which to gauge the effects of size mismatch, we compiled estimation performance statistics for each of these three dihedrals, using dihedral scattering models constructed from the actual dihedral dimensions and other primitives' scattering models constructed from the dimensions in Table 6. The results from 500 Monte Carlo runs with a 10-deg AVSD are presented in Table 11 in a format similar to that of previous sections. (For the sake of compactness, type estimation results have been abridged to a single P_{id} statistic, denoting the fraction of runs in which type was estimated correctly as a dihedral, conditional on detection.) The variation in the statistics for each of these dihedral sizes is primarily a result of the broader out-of-plane responses of shorter dihedrals, making them easier to detect.

Table 12 presents results from two experiments in which modeled dihedral dimensions are mismatched to the actual primitive dimensions. In both cases the dihedral scattering model was constructed using the square dihedral; the actual primitive, however, was either the short or tall dihedral. Comparing the results of Table 12 to those of Table 11, it is apparent that size mismatch has a relatively minor effect on performance for this example. This is an indication that a single instance of a dihedral might suffice to serve as an exemplar for the entire class of dihedrals, and, more generally, that the nominal-sizing approach used here could suffice to capture the behavior of primitives of a range of dimensions without significant degradation in performance.

6 Conclusion

The principal contribution of this work is the development of a flexible framework for the construction of 3-D target models from SAR data. The research presented here establishes a formalization for the general problem of model generation and describes a computational engine that provides a foundation for further development toward more advanced, less restrictive model generation. There are three fundamental factors to be weighed when considering any extension to this basic framework: the improvement in physical fidelity, the effect on algorithm performance, and the impact on algorithm complexity. Greater physical fidelity will generally be rewarded with better performance, but will also be attended by an increase in complexity. An extension that achieves a slight improvement in performance at the cost of a significant increase in algorithm complexity is of questionable practical value.

There are several sources of loss of fidelity in our algorithm. First is an idealized and limited set of primitive responses. Greater fidelity could be attained by broadening the set of canonical responses and expanding the primitive parameterization to accommodate noncanonical or nonideal primitives. As described before, this should be done with consideration of the achievable improvement in performance and the required increase in complexity. For instance, the analysis of Sec. 5.4 suggests that algorithm performance is relatively robust to mismatches in assumed and actual dihedral dimensions; broadening primitive parameterization to include primitive dimensions might not be prudent because it would complicate the M step of the iteration with limited gain in algorithm performance. It is likely that algorithm performance is less robust to changes in primitive geometry (e.g., deviations from 90-deg trihedral and dihedral angles[23]) and thus this area is a suitable direction for extension. One approach to accommodating noncanonical primitive responses is proposed in Ref. 12.

Another limitation in fidelity is imposed by the assumed independence of report parameters as described in Secs. 2.2

Table 12 Mismatched primitive dimensions: dihedral estimation results.

			$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^x$ error [cm]		$\hat{\theta}_i^p$ rmse [*deg*]	
dihedral	P_{det}	P_{id}	bias	stdev	$\|\text{bias}\|$	$\sqrt{\text{tr(cov)}}$	az/el	rot
short	0.990	1.000	−0.207	2.010	6.784	12.147	10.167	7.894
tall	0.872	0.988	−4.062	2.799	11.741	12.715	14.182	9.850

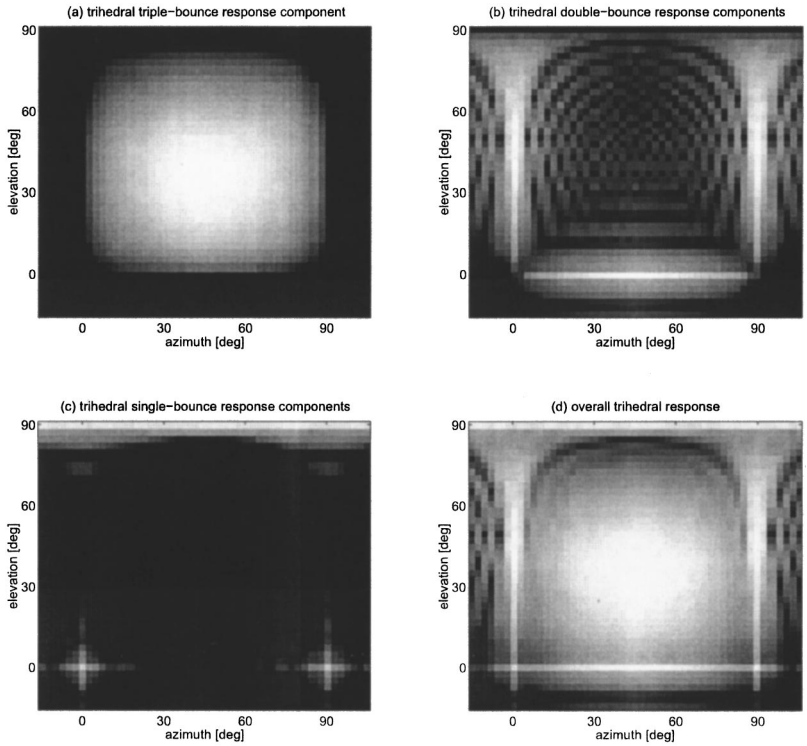


Fig. 5 Construction of trihedral response model from multiple response mechanisms.

and 2.3. For instance, the presence of single-, double-, and triple-bounce response mechanisms in the overall trihedral response suggests a correlation between $t_{k,j}$ (an even-/odd-bounce classification) and $\mathbf{x}_{k,j}$ (which projects from a different point for each response mechanism). A relaxation of Assumption 5 to model such a dependence would increase physical realism (though with unknown effect on algorithm performance) and would likely not require a significant increase in algorithm complexity: the E step could proceed largely as before, and any potential complication to the maximization would be largely mitigated by the partitioned-maximization ECM approach[16] described in Sec. 3.3. Fidelity could also be improved by utilizing a peak extractor capable of super-resolving scatterers, as discussed in Sec. 2.2, or by performing a less severe compression of the raw SAR imagery to obtain the observable features. In particular, instead of passing on a scalar amplitude for each extracted peak, we can imagine passing on a vector of continuous descriptors corresponding to a compression of the phase history associated with the peak. A useful way in which to think about such a feature vector is as the set of coefficients in a finite series approximation to the phase history: at one extreme we have a single-term series, whose sole coefficient is simply peak amplitude, while at the other extreme the coefficients would simply correspond to all of the samples of the phase history (or an invertible transform thereof). While conceptually straightforward, there are two principal challenges in developing such an extension: first, choosing a finite series form that captures information for our purposes (e.g., using bases such as those proposed by Potter and Moses[8] or by McClure and Carin[9]), and second, determining the probabilistic model for such a feature vector. In particular, the natural generalization of the Gaussian model for extracted scalar amplitude $a_{k,j}$ would be to model the vector of extracted features $\mathbf{a}_{k,j}$ as a Gaussian whose mean vector and covariance matrix would need to be estimated from data or simulations.

Another source of fidelity loss is the omission of scatterer interaction effects formalized by Assumptions 2 and 3. In particular, obstruction and multiple-primitive reflections are not modeled and are likely to have an impact on algorithm performance. Capturing multiple-primitive reflections, which arise as a result of sophisticated interactions between scatterers, would likely require fundamental modifications to the current framework and an increase in algorithm complexity due to the coupling of primitive calculations in both the E and M steps of the algorithm. Obstruction involves less complicated interactions and requires some description of how primitive detections are coupled within an image and between images. An extension that accommodates obstruction without necessitating a departure from the independent primitive-by-primitive calculations of the E and M steps is proposed in Ref. 12.

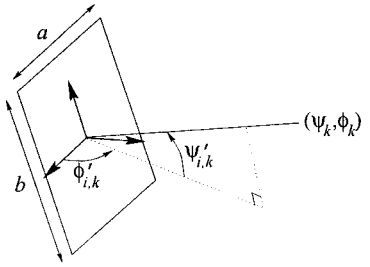


Fig. 6 Rectangular plate. The notation here is identical to that of Fig. 3.

Table 13 Physical optics RCS approximations for basic scattering mechanisms.

response mechanism	(amplitude) · (pol) · (shaping function)
rectangular plate	$\left(\frac{k_0^2 a^2 b^2}{\pi}\right) \cdot \gamma_1^{\text{pol}} \cdot [\cos\psi'_{i,k} \cos\phi'_{i,k} \operatorname{sinc}(k_0 b \sin\psi'_{i,k}) \operatorname{sinc}(k_0 a \cos\psi'_{i,k} \sin\phi'_{i,k})]^2$
dihedral double-bounce	$\left(\frac{2k_0^2 a^2 b^2}{\pi}\right) \cdot \gamma_2^{\text{pol}} \cdot \left[\sqrt{2}\cos\psi'_{i,k} \sin\left(\frac{\pi}{4} - \left\lvert\phi'_{i,k} - \frac{\pi}{4}\right\rvert\right) \operatorname{sinc}(k_0 b \sin\psi'_{i,k})\, W_{[0,\pi/2]}(\phi'_{i,k})\right]^2$
trihedral triple-bounce	$\left(\frac{3k_0^2 a^4}{\pi}\right) \cdot \gamma_3^{\text{pol}} \cdot \begin{cases} \frac{1}{3}\left(\frac{4 l_i m_i}{n_i}\right)^2 W_{[0,\pi/2]}(\phi'_{i,k}) W_{[0,\pi/2]}(\psi'_{i,k}), & m_i \leq \frac{n_i}{2}, \\ \frac{1}{3}\left[l_i\left(4 - \frac{n_i}{m_i}\right)\right]^2 W_{[0,\pi/2]}(\phi'_{i,k}) W_{[0,\pi/2]}(\psi'_{i,k}), & m_i > \frac{n_i}{2}, \end{cases}$ where $(l_i, m_i, n_i) = \text{sort}_{\min\to\max}(\sin\psi'_{i,k}, \cos\psi'_{i,k}\sin\phi'_{i,k}, \cos\psi'_{i,k}\cos\phi'_{i,k})$
cylinder shaft	$(k_0 r h^2) \cdot \gamma_4^{\text{pol}} \cdot [\cos\psi'_{i,k} \operatorname{sinc}(k_0 h \sin\psi'_{i,k})]^2$
tophat double-bounce	$(2k_0 r h^2) \cdot \gamma_5^{\text{pol}} \cdot \left[\sqrt{2}\sin\left(\frac{\pi}{4} - \left\lvert\psi'_{i,k} - \frac{\pi}{4}\right\rvert\right)\right]^2$

7 Reflector Primitive Scattering Models

As indicated in Sec. 2.1, our log-RCS response models are based on physical optics. Physical optics enables characterization of the RCSs of a wide variety of primitives as functions of the illuminating frequency, relative viewing angle $(\psi'_{i,k}, \phi'_{i,k})$, and a handful of size parameters.[5,6] The relevant size parameters for each of our four primitives are illustrated in Fig. 3. Each of these primitives embodies multiple response mechanisms. For instance, a trihedral can produce not only a triple-bounce response but also single- and double-bounce responses, and a cylinder can produce a response not only from its shaft, but also from its flat ends. Figure 5 depicts the contribution of individual response mechanisms to the overall trihedral response model: the contributing components include a triple-bounce reflection mechanism [Fig. 5(a)], three double-bounce reflection mechanisms produced by pairs of the trihedral's plates [Fig. 5(b)], and three single-bounce reflection mechanisms produced by the individual plates [Fig. 5(c)]. Although the three individual double-bounce components of Fig. 5(b) are identical except for unique Euler angle orientations, as are the three single-bounce components of Fig. 5(c), the projection of azimuth and elevation coordinates onto a rectangular array distorts each response differently. These responses are summed to yield the overall response in Fig. 5(d). Azimuth and elevation are defined as in Fig. 3. (The trihedral dimension used to generate the responses of Fig. 5 is 28.0 cm, as in Table 6.)

We construct all primitive response models from sums of five basic component response mechanisms: a rectangular plate, a dihedral double-bounce mechanism, a trihedral triple-bounce mechanism, a cylinder shaft, and a tophat double-bounce mechanism. The latter four mechanisms and their coordinate axes are clearly pictured as components of the primitives of Fig. 3; the rectangular plate is pictured in Fig. 6. The physical optics RCS approximations for the component response mechanisms can each be written as the product of three terms: a size-dependent amplitude describing the overall intensity of the response, a scaling term encompassing the effect of the given antenna polarimetry on RCS, and a shaping function describing the variation in response as a function of the relative viewing angle. Let k_0 denote the wavenumber of the illuminating frequency, let $\operatorname{sinc}(x) = \sin(x)/x$, and define a window function

$$W_{[\alpha,\beta]}(x) = \begin{cases} 1, & \alpha \leq x \leq \beta, \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

Physical optics then yields the RCS approximations given in Table 13.[5,6] The γ_i^{pol} terms in each row are polarimetric scaling coefficients.[11,24] If fully polarimetric measurements are available and the RCS is taken to be the magnitude of the polarimetric vector $[HH \ \sqrt{2}HV \ VV]$, then $\gamma_i^{\text{pol}} = 2$

Table 14 Primitive scattering model composition.

primitive	θ_i^t	θ_i^a	components of $S_{\theta_i^t}(\psi'_{i,k}, \phi'_{i,k})$	relevant dimensions
trihedral	1	$10\log_{10}\left(\frac{3k_0^2 a^4}{\pi}\right)$	one trihedral triple-bounce response three dihedral double-bounce responses three rectangular-plate responses	a
tophat	2	$10\log_{10}(2k_0 r h^2)$	one tophat double-bounce response one cylinder-shaft response	r, h
dihedral	3	$10\log_{10}\left(\frac{2k_0^2 a^2 b^2}{\pi}\right)$	one dihedral double- bounce response two rectangular-plate responses	a, b
cylinder	4	$10\log_{10}(k_0 r h^2)$	one cylinder-shaft response	r, h

for all components. If only single-polarimetric data are available, the γ_i^{pol} will take on values between 0 and 1.[11,24] The shaping functions of Table 13 are all scaled to give a maximum value of 1.

As described in Sec. 2.1, our scattering models are formed as noncoherent sums of components of Table 13 with empirically chosen nominal primitive dimensions. Recalling Eq. (3), c_{pol} is taken to be the log of the polarimetric scaling constant, $S_{\theta_i^t}$ is taken to be the log of the sum of the appropriately rotated and scaled component response shaping functions, and θ_i^a is taken to be the log of the amplitude scaling factor that ensures $S_{\theta_i^t}$ will have a maximum value of 0 dBsm. Each primitive's θ_i^a, the mechanisms that we use to construct its $S_{\theta_i^t}$, and the relevant dimensions affecting $S_{\theta_i^t}$ are given in Table 14. The assumed primitive dimensions used to construct the $S_{\theta_i^t}$ are indicated in Sec. 5 for all results presented there.

Acknowledgments

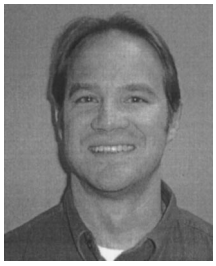
This research was supported by AFOSR grant F49620-98-1-0349, by an ODDR&E MURI through AFOSR grant F49620-96-1-0028 (subcontract GC123919NGD from Boston University), and by a National Science Foundation Graduate Research Fellowship.

References

1. J. C. Curlander, *Synthetic Aperture Radar: Systems and Signal Processing*, John Wiley and Sons, New York (1991).
2. C. V. Jakowatz, Jr., D. E. Wahl, P. H. Eichel, D. C. Ghiglia, and P. A. Thompson, *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*, Kluwer Academic Publishers, Boston (1996).
3. D. E. Dudgeon and R. T. Lacoss, "An overview of automatic target recognition," *Lincoln Lab. J.* **6**(1), 3–10 (1993).
4. E. R. Keydel and S. W. Lee, "Signature prediction for model-based automatic target recognition," *Proc. SPIE* **2757**, 306–317 (1996).
5. G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum, *Radar Cross Section Handbook*, Plenum Press, New York (1970).
6. E. F. Knott, J. F. Shaeffer, and M. T. Tuley, *Radar Cross Section*, Artech House, Norwood, MA (1993).
7. H.-C. Chiang, R. L. Moses, and L. C. Potter, "Model-based classification of radar images," *IEEE Trans. Inf. Theory* **46**(5), 1842–1854 (2000).
8. L. C. Potter and R. L. Moses, "Attributed scattering centers for SAR ATR," *IEEE Trans. Image Proc.* **5**(1), 79–91 (1997).
9. M. McClure and L. Carin, "Matching pursuits with a wave-based dictionary," *IEEE Trans. Signal Process.* **45**(12), 2912–2927 (1997).
10. E. C. G. Sudarshan and N. Mukunda, *Classical Dynamics: A Modern Perspective*, John Wiley and Sons, New York (1974).
11. *Radar Polarimetry for Geoscience Applications*, F. T. Ulaby and C. Elachi, Eds., Artech House, Norwood, MA (1990).
12. J. A. Richards, "Target model generation from multiple synthetic aperture radar images," Ph.D. thesis, M.I.T., Cambridge, MA (2001).
13. J. J. Sacchini, W. M. Steedly, and R. L. Moses, "Two-dimensional Prony modeling and parameter estimation," *IEEE Trans. Signal Process.* **41**, 3127 (1993).
14. Y. Bar-Shalom and T. E. Fortmann, *Tracking and Data Association*, Academic Press, Boston (1988).
15. T. Kurien, "Issues in the design of practical multitarget tracking algorithms," in *Multitarget-Multisensor Tracking: Advanced Applications*, Chap 3, Y. Bar-Shalom, Ed., pp. 43–83, Artech House, Norwood, MA (1990).
16. G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*, John Wiley and Sons, New York (1997).
17. A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm (with discussion)," *J. R. Stat. Soc. Ser. B. Methodol.* **39**, 1–38 (1977).
18. C. Wu, "On the convergence properties of the EM algorithm," *Ann. Stat.* **11**, 95–103 (1983).
19. M. R. Anderberg, *Cluster Analysis for Applications*, Academic Press, New York (1973).
20. M. Hazlett, D. J. Andersch, S. W. Lee, H. Ling, and C. L. Yu., "XPatch: a high-frequency electromagnetic scattering prediction code using shooting and bounding rays," in *Targets and Backgrounds: Characterization and Representation*, W. R. Watkins and D. Clement, Eds., Vol. 2469, pp. 266–275 (1995).
21. *User's Manual for XPatchT*, version 2.4, DEMACO, Inc., Champaign, IL (1991).
22. L. M. Novak, M. C. Burl, and W. W. Irving, "Optimal polarimetric processing for enhanced target detection," *IEEE Trans. Aerosp. Electron. Syst.* **29**(1), 234–244 (1993).
23. W. C. Anderson, "Consequences of nonorthogonality on the scattering properties of dihedral reflectors," *IEEE Trans. Antennas Propag.* **AP-35**(10), 1154–1159 (1987).
24. J. R. Huynen, "Phenomenoloigcal theory of radar targets," Ph.D. thesis, Delft University of Technology, The Netherlands (1970).

John A. Richards recevied a PhD in electrical engineering and computer science from MIT in 2001. He is currently a senior member of the technical staff in the Signal and Image Processing Systems Department of Sandia National Laboratories in Albuquerque. Prior to joining Sandia, he was a research assistant in the Stochastic Systems Group of the Laboratory for Information and Decision Systems at MIT. He has served as teaching assistant and recitation instructor for MIT's graduate-level course in detection, estimation, and stochastic processes. He received SB and MEng degrees in electrical engineering from MIT in 1996. He was a National Science Foundation Graduate Research Fellow and is a member of SPIE, Tau Beta Pi, and Eta Kappa Nu. His current research focuses on algorithms for automatic target recognition. His general research interests involve statistical signal and image processing, detection and estimation, and optimization.

Alan S. Willsky joined the MIT faculty in 1973 and is currently the Edwin S. Webster Professor of Electrical Engineering. He is a founder and member of the board of directors of Alphatech, Inc., and a member of the US Air Force Scientific Advisory Board. He has received several awards including the 1975 American Automatic Control Council Donald P. Eckman Award, the 1979 ASCE Alfred Noble Prize, and the 1980 IEEE Browder J. Thompson Memorial Award. He had held visiting positions in England and France and various leadership positions in the IEEE Control Systems Society (which made him a distinguished member in 1988). He has delivered numerous keynote addresses and is coauthor of the undergraduate text *Signals and Systems*. His research interests are in the development and application of advanced methods of estimation and statistical signal and image processing. Methods he had developed have been successfully applied in a variety of applications including failure detection, surveillance systems, biomedical signal and image processing, and remote sensing.



John W. Fisher III received a PhD degree in electrical and computer engineering from the University of Florida, Gainesville, in 1997. He is currently a research scientist in the Artificial Intelligence Laboratory and affiliated with the Laboratory for Information and Decision Systems at MIT. Prior to joining MIT, he was affiliated with the University of Florida, as both a faculty member and graduate student since 1987, during which time he conducted research in the areas of ultra-wideband radar for ground penetration and foliage penetration applications, radar signal processing, and automatic target recognition algorithms. His current area of research, begun as a PhD candidate in the Computational NeuroEngineering Laboratory at the University of Florida, is in information theoretic approaches to signal processing, machine learning and computer vision. He is a member of IEEE and SPIE.