Target Model Generation from Multiple Synthetic Aperture Radar Images

by

John A. Richards

S.B., Electrical Science and Engineering
Massachusetts Institute of Technology, 1996

M.Eng., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1996

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

June 2001

Signature of Author: ____________________
Department of Electrical Engineering and Computer Science
May 21, 2001

Certified by: ____________________
Alan S. Willsky
Edwin S. Webster Professor of Electrical Engineering
Thesis Supervisor

Accepted by: ____________________
Arthur C. Smith
Professor of Electrical Engineering
Chair, Committee on Graduate Students

Target Model Generation from Multiple Synthetic Aperture Radar Images

by

John A. Richards

Submitted to the Department of Electrical Engineering and Computer Science
on May 21, 2001 in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Electrical Engineering and Computer Science

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 computer-aided-design (CAD) models. Recently, efforts to generate models from a single one-dimensional radar range profile or a single two-dimensional synthetic aperture radar (SAR) image have met with some success. However, the models generated from these data sets are of limited use to most ATR systems because they are not three-dimensional. We propose a framework for generating a three-dimensional target model directly from multiple SAR images of a target obtained at arbitrary viewing angles. We cast model generation as a parametric estimation problem in which we seek a description of the target in terms of its component reflector primitives given a set of features extracted from the SAR images. We accomplish this parametric estimation in the context of data association using the expectation-maximization (EM) method. We construct an EM-based algorithm that generates a target model from features extracted from a set of SAR imagery. This algorithm adaptively selects the model order and operates without supervision. We present extensions to the basic model-generation framework and algorithm to enable explicit accommodation of various phenomena important for proper target modeling, including partial obstruction and noncanonical primitive responses. Throughout the thesis we present results demonstrating the applicability of our framework and the utility of our algorithm.

Thesis Supervisor: Alan S. Willsky
Title: Edwin S. Webster Professor of Electrical Engineering

Acknowledgments

As I find myself mere days away from bestowal of the title “Doctor of Philosophy,” I realize that I have precious little of my own philosophy to offer readers of this thesis. Early in my graduate-student career I envisioned myself capping off this document with a cleverly written acknowledgments section shimmering with pearls of wisdom. As it turns out, most of the philosophy I now possess has either been mined in the form of three-minute audio nuggets or has been gleaned from observation of the good people around me. Perhaps the best I can offer from the former source is the following: “Where you've been is good and gone, all you keep is the getting there” [106]. The philosophy from the latter source is more difficult to quote, and it does not sound as good on the guitar. It is just as valuable, however, and in these two pages I have tried to express my gratitude to those who have been kind enough to impart it.

Thanks are first due to Alan Willsky, my research supervisor. Through the last four-and-a-half years, Alan has served not only as a technical advisor but also as a teacher, mentor, counselor, friend, and fellow hard-luck Red Sox fan. He has excelled in each of these roles. Alan's intelligence and expertise are exceeded only by his compassion and understanding. He has served—and will continue to serve—as an exemplary role model, both professionally and personally.

I would also like to thank the other members of my thesis committee. Each one of them has an incredible intuitive grasp of technical issues and a mind like a steel trap. (This last metaphor is intended to evoke images like “sharp and incisive,” not “rusty and illegal in 41 states.”) John Fisher has provided numerous helpful suggestions and ideas in our many discussions over the last several years. Jeff Shapiro has been giving me excellent advice and helpful criticism since we first met nearly six years ago. Bill Irving has lent considerable insight and was instrumental in helping to formulate the basic research problem of this thesis four years ago.

During my tenure in the Stochastic Systems Group I have been surrounded by many exceptional people. Andrew Kim has been a fellow SAR guy, SPIE roommate, 6.433 and 6.252 classmate, OQE study partner, and most of all a great friend. His selfless devotion to the computer system has kept the lab running smoothly; his irrational devotion to USC and the rest of the Pac-10, despite the clear and overwhelming superiority of the Big Ten, has kept him from winning his football pools. Mike Schneider shares my unusual taste in movies and politics, two subjects we have discussed frequently.

He has been a source of helpful advice and unfailing good humor, and I owe him an apology for still having only a vague conception of what a Krylov subspace really is. Cedric Logan provided the valuable opportunity to trade a small fraction of my average weekly productivity for a large return in my average weekly enjoyment; his laid-back style, repertoire of muttered one-liners, and religious stewardship of the manifesto board all made SSG a better place. Martin Wainwright has provided many hours of enjoyable conversation and a valuable illustration of the effects of unchecked entropy in one's workplace. He also deserves recognition for tolerating many dubious attempts at Canada-related humor. (Despite years of ribbing he has not attacked any of us with the jagged, broken-off neck of a Molson's bottle.) Dewey Tucker has been a good friend and study partner despite having a schedule completely orthogonal to mine. Andy Tsai has been a model of diligence and shares the blame with me for causing an entire class of EECS graduate students to lose confidence in their calculus abilities. These and other SSG students—Terrence Ho, Austin Frakt, Ilya Pollak, Ron Dror, Alex Ihler, Junmo Kim, Erik Sudderth, and Jason Johnson—have made this lab a fun and stimulating place.

In addition to those just mentioned, two people have made the last few years at MIT enjoyable (at best) or bearable (at worst) through their companionship. Josh Winn has weathered most of the last ten years with me. We have shared coffee breaks, squash matches, beers, road trips, and above all a close friendship that has been invaluable to me. I owe him many thanks. Taylore Kelly combines a unique perspective with a refreshing openness and a rare, deep compassion for others. She has always been willing to listen and has always offered good advice. Her genuine empathy and offbeat humor have brightened my long days at MIT. I am lucky to have such a fine person as a friend.

I have received immeasurable encouragement and moral support from my family. Jim and Barb Wiley have shown genuine warmth and have been incredibly supportive of a son-in-law who has not held a real job since he married their daughter. My sister, Heather Richards, has taught me more than I have any right to have learned from a younger sibling. My parents, John and Marilyn Richards, have been a source of love and encouragement throughout my life. Their undying devotion and many sacrifices have allowed me to reach the place I now occupy. I have thought long and hard about how I might repay them, and I have come to the conclusion that the only possible way is to do for somebody else what they have done for me, somehow, sometime, somewhere down the line. Stay tuned.

Finally, I owe to my wife Becky more thanks than it is possible to express. She is a constant and limitless source of inspiration. She has shown nothing but support throughout this long endeavor, even during these last few months when she has been a thesis widow. Her unwavering love, strength, intelligence, and understanding have made the last several years a joy, even when details conspired to make individual days or weeks a misery. Through her love I have become a better person than I was when I began the research for this thesis over four years ago.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

IMAGE data plays a central role in many fields and applications. In recent years, especially as remote sensing technologies and computing resources have progressed markedly, there has been a proliferation of applications in which image data is of fundamental importance. In many fields, the collection of the image data is only one of a number of difficult technical challenges: often, the most daunting task is the intelligent processing of the collected imagery to infer something about the environment or the scene being imaged. The imagery collected by a weather satellite, or a radio telescope, or a magnetic resonance imaging system is of interest only insofar as it can be processed by a human or machine to extract information relevant to tomorrow's weather, or a distant galaxy, or a patient's diagnosis.

In many fields the primary goal of the image processing is to identify and characterize specific objects of interest, or *targets*, in the presence of significant noise and uninformative background features, or *clutter*. One such field is military reconnaissance [82]. In a typical reconnaissance scenario, the imaging system might consist of a sensor mounted on an airplane; the scene being imaged might contain a wide variety of natural features (*e.g.*, trees, grass, and rivers), fixed manmade structures (*e.g.*, buildings and roads), and vehicles (*e.g.*, trucks and tanks). In such a scenario the vehicles are often the targets of interest, and all other scene components are viewed as distracting clutter. In order to make intelligent decisions based on imagery collected in such a scenario, many questions must typically be answered: which parts of the image correspond to targets, as opposed to clutter? How many vehicles are present? Where are these vehicles located? What types of vehicles are they? Are these vehicle types indicative of friends or foes? If only a small amount of imagery is to be processed, then this processing might be accomplished by a human image analyst; if, on the other hand, there is a significant amount of imagery and time is of the essence, then it would be beneficial to employ an automated processor. Automatic target recognition (ATR) systems have been designed to process data and to identify and classify targets of interest in scenarios like these [9, 10, 31].

ATR systems have received considerable attention in military reconnaissance applications, especially in conjunction with synthetic aperture radar (SAR) imaging systems [15, 23, 51, 76, 77, 102]. The broad utility of SAR as a remote sensing tool is well known: it is effective at large distances, at day or at night, in any weather conditions,

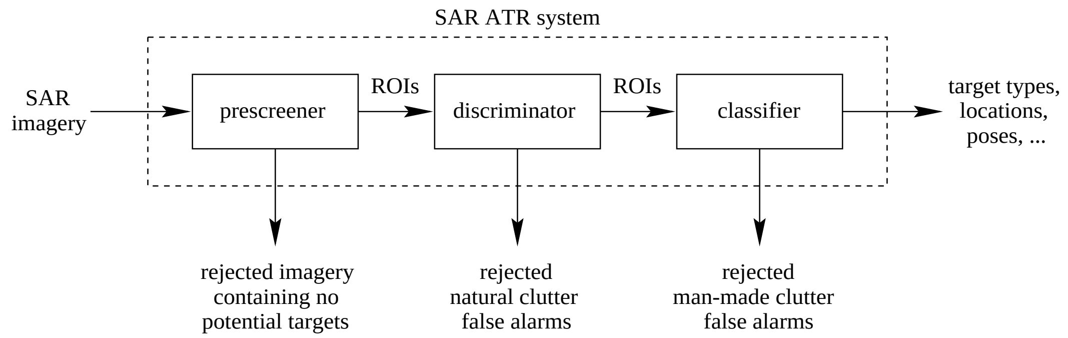


Figure 1.1. A three-stage ATR system.

and it is capable of producing high-resolution images, all while using a sensor small enough to be mounted on an airplane. Research into ATR systems operating on SAR imagery has intensified in recent years as more and more SAR imagery has become available for wide release and experimentation [73, 74].

The high-level motivation for this thesis is the generation of target models from SAR imagery to facilitate development and deployment of ATR systems. In particular, this thesis is concerned with the question of how compact, descriptive, and robust target models for ATR can be generated from multiple SAR images of a target. This introductory chapter elucidates both the context and motivation of this problem, and provides an outline for the remainder of the thesis. We begin by sketching out the context in which the problem of target model generation is important.

■ 1.1 Overview of Automatic Target Recognition

A typical ATR system for SAR is described in [81]. This system has three distinct processing stages, as depicted in Figure 1.1: a prescreener, a discriminator, and a classifier. The prescreener processes large amounts of imagery to identify a relative handful of *regions of interest* (ROIs), or small image chips potentially containing targets. The discriminator more closely examines these ROIs and rejects those that appear to contain only natural clutter, sending the rest to the classifier for even closer scrutiny. The classifier makes a final determination about each ROI, either rejecting it as uninteresting manmade clutter or identifying it as a target and reporting estimates of relevant target information such as precise location, vehicle type, pose, and other deployment conditions. The benefit of the multi-stage approach of Figure 1.1 is the tradeoff it allows between image volume and algorithmic complexity: the prescreener must process all collected imagery to screen out clutter, but can be implemented according to a very simple criterion (such as amplitude thresholding); the discriminator can more closely examine the relatively few image regions that are passed from the prescreener; the classifier can then apply sophisticated processing to the remaining image chips, which represent a very small subset of the original data set.

All ATR systems, whether implemented as in Figure 1.1 or not, employ some sort of classifier. An ATR system's classifier typically does not operate directly on collected imagery, but rather on vectors of *features*, or descriptive parameters extracted from each ROI. Features can be descriptions of extracted peaks or edges, or descriptions of segmented image regions, or basis function decompositions, or any other description of an ROI that tends to compress it into a handful of salient parameters. Classification is typically accomplished by comparing the extracted ROI feature vector to a database of target models. Depending on the structure and implementation of this database, ATR systems can be categorized as one of two types: *template-based* systems or *model-based* systems. Template-based ATR systems maintain a large database of exemplar feature vectors, or *templates*, taken from real or synthetic imagery of a library of targets, typically at a number of different poses and in a variety of deployment conditions and viewing configurations; each database template provides a classification hypothesis, and candidate targets from collected imagery are classified based on how well they match these templates. Model-based ATR systems, on the other hand, maintain a database of physical or conceptual models of targets that can be manipulated on the fly to predict feature vectors for targets in arbitrary poses, deployment conditions, and viewing configurations. Classification is achieved by intelligently refining the model selection and manipulation process until a feature-vector classification hypothesis is obtained that closely matches the extracted feature vector. (The ATR system described in [81] is template-based; a model-based ATR system is described in [112] and [36].)

Traditionally, template-based systems have found wider use than model-based systems. This is due in large part to the relative conceptual and computational complexity of a model-based system's classifier, which requires high-level reasoning, model manipulation, and feature-vector prediction in near-real-time in order to classify targets. In recent years, however, there has been a surge of interest in model-based ATR systems [36], due in large part to computational advances that make their implementation feasible. Model-based systems offer several conceptual advantages over template-based systems, including a more general framework and potentially a more compact database. Additionally, the model-based framework allows for a continuous and arbitrarily fine selection of target pose and deployment conditions in the selection of classification hypotheses, as opposed to the discrete selection imposed by the structure of the template-based framework. Unfortunately, the development and deployment of model-based ATR systems is hampered by the difficulty of creating target models to populate the ATR database. The problem of generating target models is the focus of this thesis.

1.2 Thesis Motivation and Contributions

The generation of target models to populate the target-model database is a problem that is central to the implementation of any model-based ATR system. One conceptual basis for a model-based ATR-system database is to represent targets in terms of a small set of canonical *reflector primitives* such as flat plates, cylinders, tophats, dihedrals, and trihedrals [61, 92]. This approach offers numerous benefits to a model-based ATR system. In particular, reflector primitives enable compact description of a rich class of targets in a format that lends itself directly to feature-based processing; they couple physical relevance to predictive utility, facilitating the model manipulation and component articulation required to form classification hypotheses; they allow statistical uncertainty and variability to be modeled directly in feature space; finally, they are more stable and potentially more informative than full facetization or computer-aided-design (CAD) models and potentially more straightforward to generate from SAR imagery. Facetization and CAD models [61, 118], or so-called "scattering center" models derived from facetization and CAD models, typically serve as the basis for classifier databases in current model-based ATR systems [56]. However, the very targets which are of utmost interest in ATR applications are often those for which no such detailed prior information is available.

This thesis focuses on the problem of creating a three-dimensional reflector-primitive target model from multiple SAR images of a target. It is motivated in part by recent model-based ATR research, much of which has considered the problem of extracting relevant reflector-primitive-based features from single view radar measurements in the form of range profiles [16, 45, 70, 85] or SAR images [86, 93]. These efforts have been successful in demonstrating that information pertaining to the reflector-primitive composition of a target can be robustly and reliably extracted from radar data. Our work builds on this conceptual legacy and considers the problem of how reflector-primitive-based information can be robustly and reliably extracted when the available radar data is not limited to a single viewing angle, but instead is spread over multiple viewing angles across the viewsphere.

The fundamental contribution of this thesis is the development and demonstration of a systematic formalism and flexible framework for the generation of three-dimensional models from 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 [56, 92]. 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

Figure 1.2. Two views of SLICY.

operate in this manner, we 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 data extracted from the individual images or phase histories. These compressed representations are then fused in order 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. In order to clearly present 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.

A motivating example for the research presented here is the target depicted in Figure 1.2. This target, known as SLICY (an acronym for Sandia Laboratories Implementation of Cylinders), although simpler than many manmade targets such as trucks or tanks, is sophisticated enough to exhibit many or most of the complicated real-world effects observed in more complex targets. SLICY is a target that exhibits noncanonical primitive responses, cavity responses, primitive coupling in the form of multiple-bounce reflections, and is self-occluding from many viewing angles. A motivating question for much of the work in this thesis is the following: can we develop a model generation framework that is sophisticated enough to deal with a target like SLICY?

■ 1.3 Thesis Organization and Main Contributions

The remainder of this thesis is organized as follows.

Chapter 2, Background and Previous Work

We begin by describing the basic physical principles on which radar systems are based. We describe how radar systems achieve resolution in the range and cross-range directions. This leads to a discussion of SAR, in which high resolution is achieved by clever processing of sequential radar returns. We introduce several concepts that are instrumental in the understanding and interpretation of SAR images, and that aid in the construction of a measurement model in Chapter 3. We consider the propagation phenomenology that dictates the appearance of reflector primitives in SAR images and describe how several simplifying assumptions enable concise description of the appearance of such primitives; the assumptions and models underlying the methods of Geometrical Optics (GO) and Physical Optics (PO) [61, 92] are given particular attention due to their extensive utilization in later chapters. We conclude the chapter with a description of recent work in which GO, PO, and related models have been used to extract reflector-primitive-based features from 1-D and 2-D radar measurements in the form of range profiles and SAR images.

Chapter 3, Precise Problem Formulation

In this chapter we formalize the problem statement and construct a measurement model based on the considerations of Chapter 2. We propose a framework in which the model generation problem is broken into two stages: first, the extraction of features from SAR images, and second, the fusion of these features across images to yield a 3-D target model. The benefits of this two-stage approach are numerous: it allows a great deal of flexibility in both the description of salient features in individual images and the description of the target, it transforms an otherwise ill-posed inverse problem into a well-posed correspondence problem, it casts the model generation problem into a realm in which data association techniques can be applied, and it greatly facilitates the construction of a tractable measurement model. We discuss various choices of target parameterizations and feature-extraction techniques, and for purposes of illustration select specific instances of each that are used extensively throughout the remainder of the thesis. The physical considerations of Chapter 2 are utilized in conjunction with several simplifying assumptions to produce a measurement model relating the observed SAR image features to the parametric description of the target which is to be estimated. This measurement model illuminates the true nature of the model generation problem in our framework, in which any SAR image feature has associated with it two fundamental levels of uncertainty: first, the very identity of that feature (*i.e.*, its unknown correspondence to the set of target primitives) and second, the stochastic nature of the feature parameters conditional on that unknown identity. We discuss the validity of our assumptions and allude to relaxations of these assumptions that will enable the accommodation of sophisticated real-world effects such as those described in Chapters 6 and 7.

Chapter 4, An Expectation-Maximization Approach to Model Generation

In this chapter we present an algorithm to generate 3-D target models from extracted SAR features in the framework constructed in Chapter 3. This algorithm is based on the Expectation-Maximization (EM) method [71, 75], an iterative technique for arriving at a maximum likelihood estimate of parameters in data association problems, and can conceptually accommodate any choice of target parameterization and feature extraction that can be cast in the framework of Chapter 3. We describe the basic form of the core of the algorithm (*i.e.*, the form of the expectation and maximization steps of the EM method) for the measurement model constructed in Chapter 3, and consider various practical implementation issues. The conceptual and computational benefits of the assumptions of Chapter 3 are clearly indicated; additional assumptions are made in order to ease the computational burden of the expectation and maximization steps. We describe a technique for initializing the iteration and for modifying the model order as the iteration progresses. The initialization uses an agglomerative clustering technique to produce a rough guess at the unknown identities of all features, which are then used to produce an estimate of the number of primitives and their descriptive parameters. The model-order selection adaptation operates in conjunction with the initialization, which is biased toward an overestimation of model order, to gradually correct this overfit as the EM iteration progresses. The result is a data-adaptive algorithm that arrives at a full parametric reflector-primitive target description (including model order) with no supervision.

Chapter 5, Experimental Results from the Basic Algorithm

This chapter contains experimental results that illustrate the utility of the framework and algorithm constructed in Chapters 3 and 4. We describe the experimental setup used in this and later chapters and illustrate results demonstrating the algorithm's performance on single-primitive targets in typical clutter conditions, corresponding essentially to the simplest data association problem with which the algorithm could be faced. These results serve to demonstrate the differences in observability between the basic primitive types and to establish benchmarks against which results in later sections and chapters can be gauged. We analyze these results and illustrate the various fundamental limitations that dictate the detectability of the several primitive types under consideration and impact the quality of their parameter estimates. We demonstrate that the algorithm's detection performance for each primitive type is near theoretical bounds. We present results for more sophisticated targets which present more significant data association problems, and demonstrate that the algorithm is successful at solving these correspondence problems. We conclude this chapter by analyzing the effects of a mismatch between modeled and actual primitive dimensions on algorithm performance.

Chapter 6, Dealing with Obstruction

In this chapter we examine one shortcoming of the framework constructed in Chapters 3 and 4, namely, its inability to accommodate explicitly the effects of partial obstruction of primitives. We demonstrate the detrimental effect of obstruction on algorithm performance by presenting results similar to those of Chapter 5 but with partially obstructed primitives. We describe several conceptual approaches to accommodating obstruction by expanding the framework of Chapter 3 and modifying the algorithm of Chapter 4. The conceptually optimal method of incorporating obstruction effects is shown to be impractical, and we propose a simplified approach that is a natural extension to the existing framework and does not significantly increase the computational requirements of the algorithm. The utility of this approach is demonstrated with several examples.

Chapter 7, Dealing with Noncanonicity

This chapter begins with an analysis of real-world effects that make many primitives exhibit noncanonical responses. We examine SLICY in detail and describe its various response mechanisms, which illustrate several forms of noncanonicity. We demonstrate the inability of the algorithm constructed in previous chapters to accommodate noncanonical responses properly by presenting its performance on SLICY. We describe different conceptual approaches to dealing with noncanonicity, and propose an approach that is conceptually broad enough to accommodate a wide range of noncanonicities. Concentrating on one class of noncanonicities involving perturbations of the geometrical structure of primitives, we implement this approach and demonstrate its performance on SLICY. The modified algorithm robustly and reliably generates a realistic and descriptive model of SLICY. We conclude the chapter by describing a method for distinguishing between estimated target model components that correspond to properly modeled responses and those that correspond to improperly modeled noncanonical response mechanisms.

Chapter 8, Contributions and Suggestions for Future Research

In the final chapter we summarize the major contributions of the thesis and describe several possible directions for further research. We suggest several straightforward extensions to the framework and algorithm constructed in previous chapters and conclude by discussing a handful of deeper and more fundamental extensions.

Appendices

In addition to three appendices containing supporting material for the previously mentioned chapters, Appendix D presents a summary of the notation used throughout the thesis and Appendix E provides a glossary of acronyms.

Chapter 2

Background and Previous Work

RADAR has found wide use in scientific, commercial, and military applications for a myriad of purposes. Its ubiquitousness stems from its many advantages in practical sensing scenarios: radar is a non-invasive, remote, all-weather imaging technique capable of operating at night or during the day. Radar systems are able to detect targets and produce high-resolution images from distances of many miles.

Radar relies on the facts that objects reflect radio waves, and that different kinds of objects reflect radio waves differently. Although the concept of the detection of targets by radio waves dates from the early 1900s [99], radar was not extensively developed and did not achieve widespread use until World War II, when its utility as a military tool was first exploited.[1] The years immediately following the war saw an explosion of radar technology and widespread research into target detection and classification by radar. Radar development in this era was characterized by gradually improving range resolution but generally inadequate azimuthal, or cross-range, resolution. In the 1950s, Carl Wiley, an engineer for the Goodyear Aircraft Corporation, realized that motion of the radio antenna platform could be used to enhance azimuthal resolution, and SAR was born [111]. This chapter presents some basic radar and SAR concepts, describes relevant phenomenological models for the interaction of objects with radio waves, and examines several recent efforts to extract target information from radar measurements.

■ 2.1 Radar Fundamentals

Radar measurements are made by emitting radiation and "listening" for reflections. More precisely, a radar system uses a transmitting antenna to radiate energy in periodic bursts. A receiving antenna detects the reflected signal by measuring the electric field strength as a function of time. A *monostatic* radar system uses a single antenna for both transmission and reception and thus measures *backscattered* radiation. (In practice, most radar systems are monostatic; *bistatic* radar systems employ distinct antennas for transmission and reception, and *multistatic* radar systems have a single transmitting antenna but multiple receiving antennas.) The delay between the transmission of a pulse and the reception of its reflection yields information about target

[1] An excellent popular account of the development of radar during World War II (and of the development of numerous radar applications and spinoff technologies after the war) is [13].

range; the Doppler shift of the reflection provides velocity information[2]; the strength (equivalently, attenuation) and polarization of the return signal yield information about target structure and composition.

2.1.1 Radar Cross Section

Historically, the strength of the radar return signal is considered a measurement of a target's *radar cross section* (RCS), defined as

$$\sigma = \lim_{R \to \infty} \left(4\pi R^2\right)^2 \frac{|\mathcal{E}^s|^2}{|\mathcal{E}^t|^2}, \tag{2.1}$$

where R is the distance between the antenna and target, $\mathcal{E}^t$ is the transmitted field strength, and $\mathcal{E}^s$ is the return (*i.e.*, backscattered) field strength as measured at the radar platform [62]. The quantity σ has units of area; it can be interpreted as the projected area of an isotropically scattering, perfectly conducting sphere that would backscatter the same fraction of power as the target at the given viewing angle. The limit in R is taken to remove any near-field effects such as wavefront curvature or target-antenna coupling. RCS is often specified in dBsm units, shorthand notation for square meters on a decibel scale; σ is occasionally normalized and made dimensionless by dividing by the squared wavelength of the transmitted field. Note that RCS encapsulates only the magnitude of a response, and neglects all phase information.

The quantity σ in (2.1) is idealized in the sense that it represents a limit as R approaches infinity; an actual radar measurement is, of course, taken at a finite distance and thus is of the form

$$\sigma_0 = \left(4\pi R_0^2\right)^2 \frac{|\mathcal{E}_0^s|^2}{|\mathcal{E}_0^t|^2}, \tag{2.2}$$

where $|\mathcal{E}_0^s|^2$ and $|\mathcal{E}_0^t|^2$ are measured at the radar platform, and R_0 is a range measurement that can be obtained from the delay τ_0 between transmission of a pulse and reception of its reflection, *i.e.*,

$$R_0 = \frac{c\tau_0}{2}, \tag{2.3}$$

or by more sophisticated processing (as described in Section 2.1.2). Note that the measurements of $|\mathcal{E}_0^s|^2$ will depend on the polarization of the transmitted field and the polarization of the receiving antenna; to capture these dependencies, (2.2) may be modified as follows:

$$\sigma_{tr} = \left(4\pi R_0^2\right)^2 \frac{|\mathbf{e}_r \cdot \boldsymbol{\mathcal{E}}_0^s|^2}{\|\boldsymbol{\mathcal{E}}_0^t\|^2}, \tag{2.4}$$

[2]Most modern radar systems use sophisticated signal processing techniques such as deramp compression [51], described in Section 2.1.2; most such techniques assume a stationary target and utilize Doppler-shift information to enhance range resolution, with the result that relative range and velocity information are not directly separable.

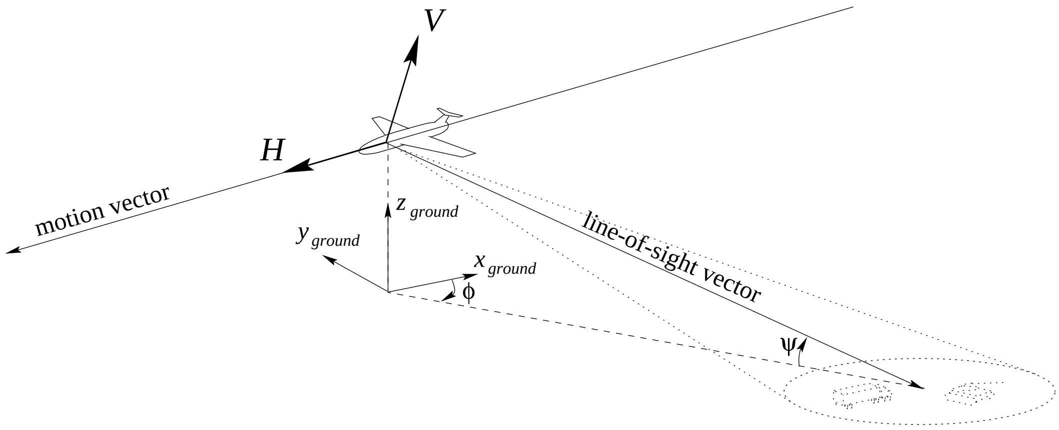


Figure 2.1. Antenna polarization and imaging geometry.

where $\boldsymbol{\mathcal{E}}_0^s$ and $\boldsymbol{\mathcal{E}}_0^t$ are now vector fields and $\mathbf{e}_r$ is a unit vector aligned with the receiving antenna polarization [38,62]. In practice the tr subscript on σ is often specified as HH, HV, VH, or VV, where H stands for "horizontal" polarization and V for "vertical" polarization according to the geometry depicted in Figure 2.1; note that H and V are defined in terms of the platform motion vector and the line-of-sight vector (which is defined in turn by the azimuth ϕ and depression ψ of the target with respect to the radar platform). Occasionally, left-handed and right-handed circularly polarized fields are also used to obtain radar measurements, in which case (2.4) still holds, but with complex field and unit vectors.[3]

A single RCS measurement characterizes only the return field strength for a particular transmit-receive polarization pair for a particular delay τ_0 (and thus for a particular range R_0). Multiple measurements of RCS from the same target are often presented in the format of a *range profile*, which is a vector of σ measurements (and often, associated phase information) at a fixed azimuth and depression, as a function of range. If range profiles at closely spaced azimuths are available, these may be concatenated to produce an image. Such images are often treated as visual data, although strictly speaking they represent measurements of the ratio of incident field strength to reflected field strength at an array of range-azimuth locations.

We can further quantify the nature of RCS measurements of a scene by introducing the concept of *complex reflectivity density* [77], which we shall denote as

$$g(x,y) = m(x,y)e^{j\rho(x,y)}, \tag{2.5}$$

where $m(x,y)$ and $\rho(x,y)$ are both real-valued functions representing the magnitude and phase of $g(x,y)$, respectively.[4] Complex reflectivity density has the following physical

[3] In fact, the above equations and definitions accommodate radar measurements taken at arbitrary elliptical polarizations, although in practice linear and circular polarizations are used almost exclusively.

[4] This formulation follows that of [77] and assumes a 2-D target; this simplifies analysis by precluding

interpretation: a target patch of area $dx\,dy$ centered at ground location (x, y) will reflect the transmitted signal with power attenuated by the factor $m(x, y)\,dx\,dy$ and with an intrinsic phase shift (*i.e.*, not attributable to the propagation delay) of $\rho(x, y)$. (In general, $m(x, y)$ and $\rho(x, y)$ are functions of viewing angle, incident signal frequency, and polarization, although this dependence has been suppressed in the notation of the previous equations.) In this formulation, RCS measurements represent the magnitude of the integration of $g(x, y)$ over the radar footprint, *i.e.*, the region illuminated by the radar. Clearly, if the radar system can be designed to illuminate smaller regions, or equivalently if returns can be processed to simulate those that would be observed from such a system, resolution will improve. We now turn our attention to the problem of how modern radar systems achieve resolution in both the range and *cross-range* (*i.e.*, azimuthal) directions.

2.1.2 Range Resolution

Range resolution of a radar system is usually defined as the minimum range separation between two point scatterers required to produce two distinct peaks in the radar return [109]. Range resolution depends on the choice of transmit pulse and on the processing of the return. Classically, range resolution was determined by achievable bandwidth and power constraints. In modern radar systems, a technique known as *deramp compression* is often used to enhance range resolution [23, 51]. Rather than relying on narrow transmit pulses and matched filters to facilitate return processing, deramp compression radars transmit a linear frequency-modulation (FM) chirp pulse which can be expressed in complex form as

$$b_c(t) = e^{j2\pi(f_c t + \frac{\eta_c}{2}t^2)} W_{[-\frac{T}{2},\frac{T}{2}]}(t), \tag{2.6}$$

where f_c is the *carrier* or *center frequency* of the pulse (*i.e.*, the instantaneous frequency at time $t = 0$), η_c is known as the *chirp rate* of the pulse, and $W_{[u,v]}(t)$ is a window function satisfying

$$W_{[u,v]}(t) = \begin{cases} 1, & u \le t \le v, \\ 0, & \text{otherwise}, \end{cases} \tag{2.7}$$

so that the total duration of the chirp of (2.6) is T. Figure 2.2 depicts the real part of a chirp pulse generated according to (2.6).

Let us define $g_\phi(u)$ to be the coherent projection of the complex reflectivity density function $g(x, y)$ onto the axis defined by the target azimuth ϕ, as in Figure 2.3.[5] If a linear FM chirp pulse is transmitted and the return signal is appropriately demodulated and low-pass filtered (details are given in [23, 51]), the resulting *deramp compressed*

---

the possibility of target self-occlusion. The 3-D case is considered in [50].

[5]The ensuing analysis relies on the assumptions that the target lies in the ground plane, that knowledge of $g(x, y)$ and the transmit pulse is sufficient to describe the radar response, and that the depression angle ψ is zero.

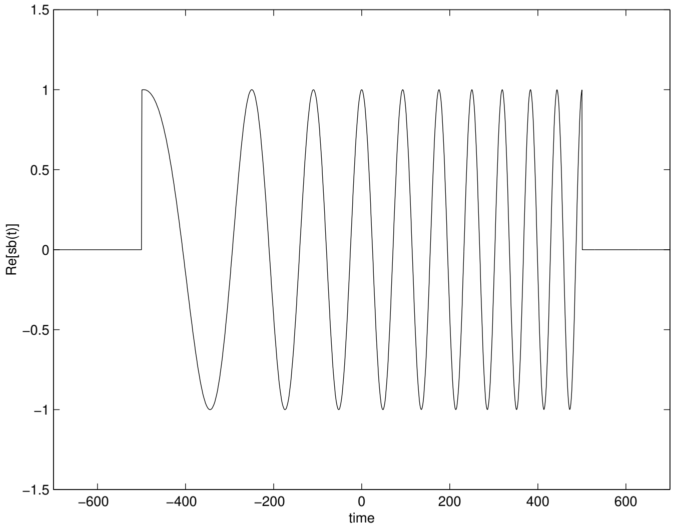


Figure 2.2. Linear FM chirp pulse.

signal (denoted $c_\phi(t)$) is approximately equal to the Fourier transform of $g_\phi(u)$ over the frequency range

$$F_1 \leq t \leq F_2, \tag{2.8}$$

where F_1 and F_2 are defined as

$$\begin{aligned} F_1 &= \frac{2}{c}\left[f_c - \eta_c \frac{T}{2}\right], \\ F_2 &= \frac{2}{c}\left[f_c + \eta_c \frac{T}{2}\right]. \end{aligned} \tag{2.9}$$

That is, $c_\phi(t)$ is approximately an interval of the Fourier transform of a projection of $g(x, y)$. Assuming at this point that we are interested only in obtaining a range profile at azimuth ϕ, we might proceed by defining

$$\hat{c}_\phi(t) = \begin{cases} c_\phi(t), & F_1 \leq t \leq F_2, \\ 0, & \text{otherwise}, \end{cases} \tag{2.10}$$

such that

$$\hat{c}_\phi(t) = \mathcal{F}\left[g_\phi(u)\right] \cdot W_{[F_1,F_2]}(t), \tag{2.11}$$

where $\mathcal{F}$ is the Fourier transform operator. Then we can obtain $\hat{g}_\phi(u)$, an imperfect representation of the projected complex reflectivity density, by taking the inverse Fourier

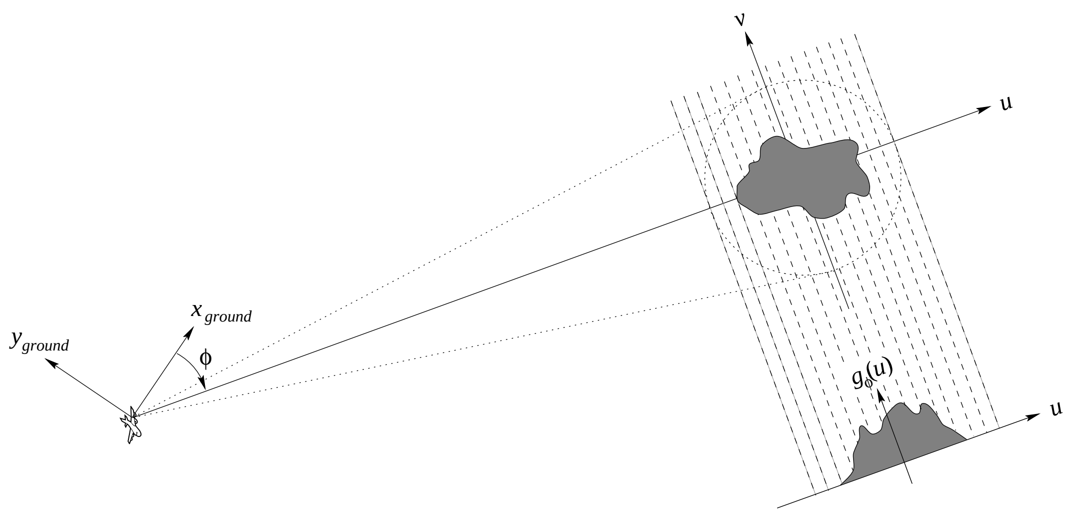


Figure 2.3. Projection of reflectivity density function.

transform of $\hat{c}_\phi(t)$. It is apparent from the above equation that this reconstruction will be a scaled version of

$$\hat{g}_\phi(u) = g_\phi(u) * \frac{2\eta_c T}{c} \operatorname{sinc}\left(\frac{2\eta_c T}{c}\pi u\right). \tag{2.12}$$

Thus the approximation $\hat{g}_\phi(u)$ is simply the convolution of the true projection $g_\phi(u)$ with a sinc function. Furthermore, (2.12) clearly indicates that the greater the duration of the transmit chirp pulse, the better the approximation $\hat{g}_\phi(u)$ will be. In fact, the range resolution for a system using deramp compression to process returns from a linear chirp transmit pulse can be shown to be [23]

$$\delta_r \approx \frac{c}{\eta_c T}. \tag{2.13}$$

This implies that we can improve resolution by increasing the duration of the transmit chirp pulse. In practice, deramp compression systems are used to achieve range resolutions on the order of 0.1 m and below [23].

■ 2.1.3 Cross-Range Resolution

As we have just seen, excellent range resolution can be achieved by choosing an appropriate transmit signal. Cross-range resolution is more difficult to achieve. Cross-range (*i.e.*, azimuthal) resolution is defined analogously to range resolution: it is the minimum cross-range separation between two point scatterers required to produce two distinct peaks in a sequence of radar returns obtained as ϕ varies. Given a specific transmit pulse, $c_\phi(t)$ is a function only of the projection of $g(x, y)$ onto the range axis. Thus, scenes with identical range projections will produce identical returns. This implies that barring clever processing, cross-range resolution is determined by the size of the region illuminated by the radar. Specifically, a scatterer's azimuthal location can be pinpointed only to the coarseness of the radar footprint. It would initially seem that the easiest way to achieve cross-range resolution, then, is to diminish the width of the radar beam.

It can be shown [15] that the cross-range resolution of the system described thus far is

$$\delta_{xr} \approx \frac{cR_0}{f_c D}, \tag{2.14}$$

where D is the antenna aperture, essentially the largest dimension of the antenna array. This equation reveals two fundamental inadequacies. First of all, it is undesirable to have resolution dependent on range, since the range of the targets being imaged will in general not be known *a priori*. To demonstrate the second, more serious problem, let us examine a typical radar imaging scenario. The quantity c is constant; physical and practical constraints make it impossible to increase f_c significantly; the use of radar as a remote sensing tool often precludes the reduction of R_0. A typical radar scenario might include a radar operating at $f_c = 10$ GHz, imaging from a distance of $R_0 = 10$ km. Suppose the application demands a cross-range resolution of 0.3 m, an easily achievable range resolution. Then (2.14) requires that $D = 1000$ m! Clearly this is much too large an antenna to be carried on board an aircraft.

Classical radar techniques require an impractically large antenna to achieve acceptable cross-range resolution. The maximum feasible antenna array size D is determined by the platform housing the radar system; the above example indicates that cross-range resolution comparable to range resolution can be achieved only if the radar is ground- or ship-based, since clearly no truck or aircraft could house an antenna array hundreds of meters long. Cross-range resolution is not achievable by classical means, and instead must be gleaned from signal processing. This is the victory of SAR.

■ 2.2 SAR Fundamentals

Figure 2.4 is a SAR image of a tank.[6] This image represents an array of backscattered RCS measurements taken at 16384 (128×128) range/cross-range pairs and presented in log-magnitude form.[7] The vertical axis of the image corresponds to range and the horizontal axis corresponds to cross-range, both relative to the scene center. The pixel resolution in both directions is approximately 0.3 m. This is an example of how SAR can achieve excellent cross-range resolution.[8]

[6] This image is T-72 chip hb03829.015 (17° depression, 296.8° azimuth) from the MSTAR (moving and stationary target acquisition and recognition) public targets data set [73].

[7] Although SAR is a coherent imaging modality offering phase as well as magnitude information, SAR images are typically presented in magnitude form only.

[8] The grainy appearance of the image is due to a SAR imaging phenomenon known as *speckle*, described below.

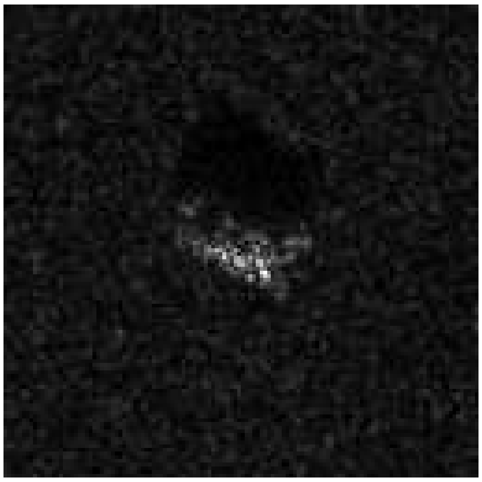


Figure 2.4. SAR image of a tank.

SAR systems achieve cross-range resolution by (as the acronym implies) synthesizing an aperture larger than the physical antenna aperture. This is accomplished by coherently processing radar returns generated at multiple points along the flight path. There are two common modes of SAR operation: *stripmap* mode and *spotlight* mode. Stripmap-mode SAR, depicted in Figure 2.5, is so named because the aircraft maintains a linear flight path while operating a sideways-looking antenna [23,76]. In this way the radar footprint sweeps out a strip on the ground. In spotlight-mode SAR, depicted in Figure 2.6, the aircraft maintains a linear flight path while continuously steering the antenna to illuminate a single patch on the ground [15, 51, 77]. In both stripmap and spotlight mode, the multiple returns from a target collected at different points along the flight path are processed to yield improved azimuthal resolution. This thesis will consider SAR systems operating in spotlight mode.

Many of the recent spotlight-mode SAR references (such as [51]) approach the subject using the tomographic formulation first proposed by Munson, O'Brien, and Jenkins in 1983 [77]. The objective of tomography is to reconstruct multi-dimensional data by processing multiple projections of that data [79, 96]. The crux of tomography is the celebrated Projection-Slice Theorem, which equates the 1-D Fourier transform of a projection of an image to a slice of the 2-D Fourier transform of the image [65]. Specifically, given data $h(x, y)$ with 2-D Fourier transform $H(f_x, f_y)$, define $h_\phi(u)$ to be the projection of $h(x, y)$ onto the axis running through x-y origin at angle ϕ, as in Figure 2.7. That is,

$$h_\phi(u) = \int_{-\infty}^{+\infty} h(u\cos\phi + v\sin\phi,\ -u\sin\phi + v\cos\phi)\, dv. \tag{2.15}$$

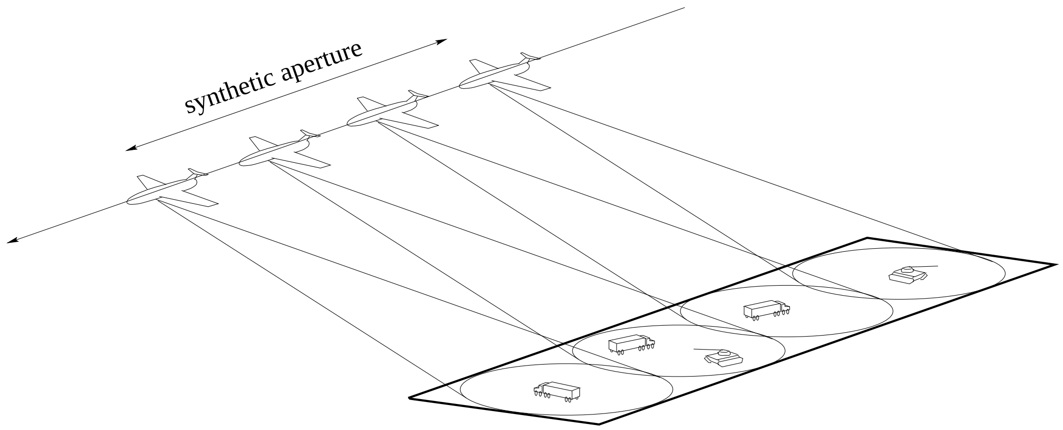


Figure 2.5. Stripmap-mode SAR.

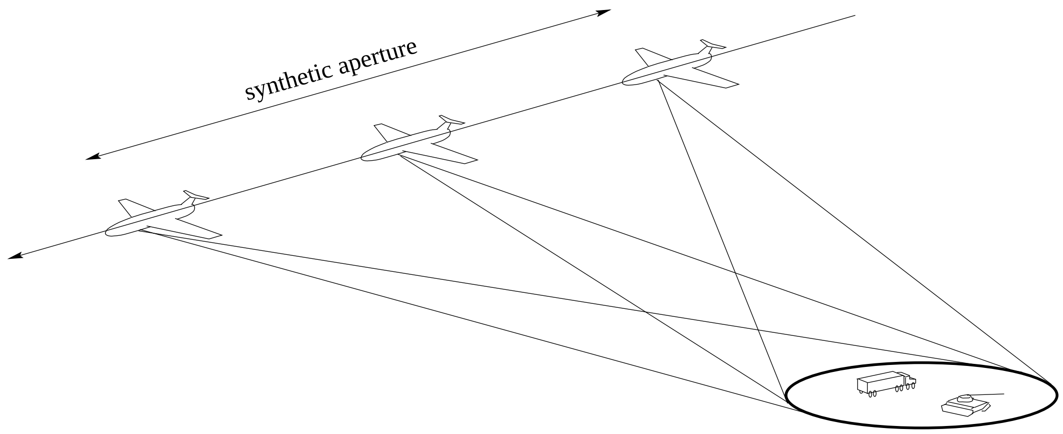


Figure 2.6. Spotlight-mode SAR.

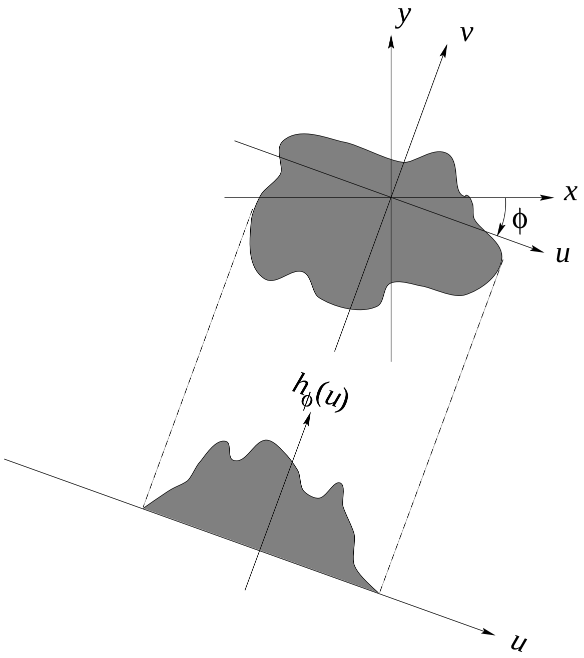


Figure 2.7. Projection from Cartesian coordinates onto an axis.

Then the Projection-Slice Theorem states

$$\int_{-\infty}^{+\infty} h_\phi(u) e^{-j2\pi f_u u}\, du = H(f_u \cos\phi, -f_u \sin\phi). \tag{2.16}$$

If a full set of projections $h_\phi(u)$ at all $0 \leq \phi < 2\pi$ is available, then the full plane of $H(f_x, f_y)$ is known and $h(x, y)$ can be reconstructed exactly.

Recall that deramp compression processing a linear chirp return yields a segment of the Fourier transform of a projection $g_\phi(u)$. That is, each return is equivalent to a portion of a slice of $G(f_x, f_y)$, the Fourier transform of $g(x, y)$. In a spotlight-mode system, returns generated at different points along the flightpath will correspond to projections at multiple angles. In this way a region of the Fourier transform of the target scene's complex reflectivity density may be constructed. The apparent shortcoming of this technique, however, is that the Projection-Slice Theorem states that the entire plane of Fourier data is required to reconstruct $g(x, y)$ exactly. The above approach will generate only a swath of Fourier data. In practice, it is often infeasible or impossible to acquire returns outside of an aperture of several degrees. This, coupled with the fact that the processed returns yield a reconstruction of the Fourier data only in a bandlimited region, means that the segment of Fourier data acquired will be very small indeed. If the range of viewing angles is small, the set of returns acquired will represent

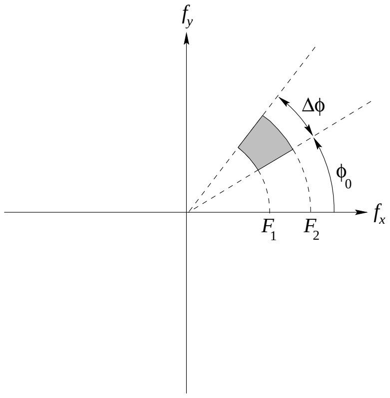


Figure 2.8. Annular segment of Fourier data.

only a roughly annular segment of Fourier data, approximately covering the region

$$\begin{aligned} F_1 \leq f_r \leq F_2, \\ \phi_0 \leq f_\theta \leq \phi_0 + \Delta\phi, \end{aligned} \tag{2.17}$$

where F_1 and F_2 are as defined in (2.9), ϕ_0 and $\Delta\phi$ are parameters defined by the viewing geometry ($\Delta\phi$ is typically no larger than a few degrees), and f_r and f_θ represent polar coordinate locations in the Fourier plane. (Figure 2.8 depicts an annular region of the Fourier plane in terms of (2.17).)

To proceed from here, one might imagine attempting to reconstruct $g(x, y)$ in the same way that $g_\phi(u)$ was reconstructed—namely, by zero-padding the return and taking an inverse Fourier transform. Let us define

$$\hat{G}(f_r, f_\theta) = \begin{cases} G(f_r, f_\theta), & F_1 \leq f_r \leq F_2 \text{ and } \phi_0 \leq f_\theta \leq \phi_0 + \Delta\phi, \\ 0, & \text{otherwise}, \end{cases} \tag{2.18}$$

and then let $\hat{g}(x, y)$ be the inverse Fourier transform of $\hat{G}(f_r, f_\theta)$. It can be shown [23] that given the bounds in (2.17), this technique achieves a cross-range resolution of

$$\delta_{xr} \approx \frac{c}{2 f_c \sin(\Delta\phi)}, \tag{2.19}$$

which means that the 10-GHz system described in Section 2.1.3 can achieve a cross-range resolution of 0.3 m with less than three degrees of azimuthal variation! This is well within the realm of feasible synthetic aperture sizes in most remote sensing applications. Another benefit is that, unlike the classical cross-range resolution result (2.14), SAR cross-range resolution does not depend on range. *A priori*, however, there is no reason to expect $\hat{g}(x, y)$, which has been constructed from an extremely bandlimited region of the Fourier plane, to resemble $g(x, y)$ at all. It seems that $\hat{g}(x, y)$ would be lacking all image characteristics not having frequencies present in the Fourier region defined by (2.17) and depicted in Figure 2.8.

As it turns out, if the seemingly naive prescription of (2.18) is followed, an acceptable image will be generated. In fact, this is how SAR images, including the one shown in Figure 2.4, are produced. In 1984, Munson and Sanz [78] suggested a theoretical basis for this phenomenon and fortified the suggestion with experimental evidence. Their argument relies on the fact that SAR is a coherent imaging modality. Very briefly, it proceeds as follows: the intrinsic phase term $\rho(x, y)$ in (2.5) is determined largely by microscopic surface phenomena at the air/target boundary [77, 104]. The microscopic characteristics of even the "smoothest" target surface are essentially random and uncorrelated except at microscopic distances. This means that macroscopic, non-overlapping measurements of $e^{j\rho(x,y)}$ are essentially uncorrelated, with energy distributed throughout the Fourier plane. Because $g(x, y)$ is a product of $m(x, y)$ and $e^{j\rho(x,y)}$, the wideband presence of the phase term modulates information about $m(x, y)$ throughout the Fourier plane. Thus, any small segment of Fourier data will contain information about $m(x, y)$ that might be recoverable.[9] (A complete description of the admittedly imprecise and incomplete argument presented here is undertaken in [78].)

Incidentally, the above description of the microscopic nature of target surfaces also offers an explanation for the SAR phenomenon known as *speckle*. (Speckle is the descriptive term for the grainy appearance of Figure 2.4 and all other SAR images.) Speckle is the visual manifestation of the randomness of the microscopic target textures. The SAR imaging process produces measurements of complex reflectivity density integrated over each pixel region. Due to the random phase fluctuations described above, some pixels will encompass microscopic scatterers that interfere constructively, while other pixels will encompass microscopic scatterers that interfere destructively. Speckle is thus an artifact of the coherent SAR imaging process.

We have now presented a basic summary of how radar and SAR work. Before proceeding, we note several things. First of all, the analysis presented in this section has been for continuous-time radar returns; in practice, the radar return $c_\phi(t)$ of (2.13) is sampled to yield a sequence of *stepped frequency* measurements $c_\phi[n]$, so named because of the Fourier-transform relationship between the range profile and $c_\phi(t)$. Consequently, SAR images are formed not from continuous $G(f_x, f_y)$ data but rather from samples of this data compiled from multiple $c_\phi[n]$ returns at different aspect angles. The array of sampled $G(f_x, f_y)$ data, or its interpolation onto a regularly spaced rectangular array on Cartesian axes, is known as a *phase history*. SAR images are formed as discrete Fourier transforms (DFTs) of phase histories.

We also note that while SAR achieves cross-range resolution unattainable by classical means, it is still able to resolve scatterers in only two dimensions. Any scatterers equidistant from the antenna in both range and cross-range will produce a superposed coherent response in the return. These contours of equidistance are approximately straight lines orthogonal to the so-called *slant plane* defined by the platform motion vector and the line-of-sight vector and depicted in Figure 2.9. Intuitively, this means

[9] According to this argument there is no reason to expect the phase of $\hat{g}(x, y)$ to conform to $\rho(x, y)$ at all, but this is unimportant if we view SAR images as measurements only of the magnitude of $g(x, y)$.

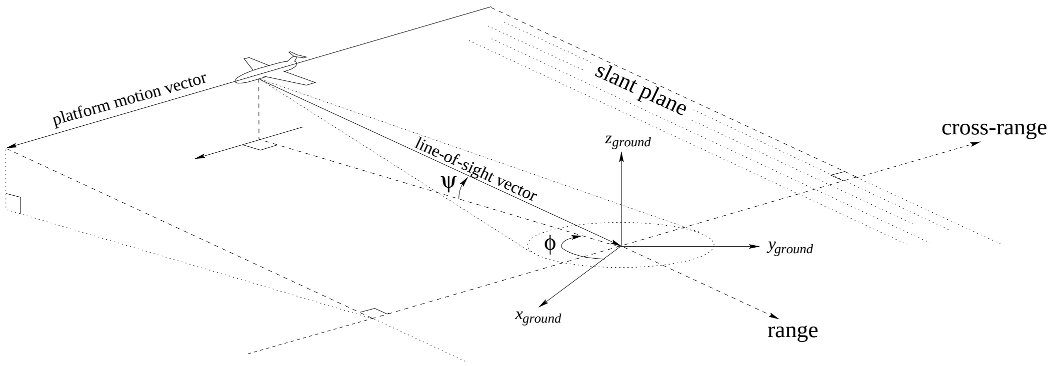


Figure 2.9. The slant plane. The slant plane is defined in terms of the platform motion vector and the line-of-sight vector. If the angle between these vectors' ground-plane components—*i.e.*, the SAR image *squint angle*—is 90° as it is here, and if the platform motion vector is parallel to the ground plane, then the slant plane is uniquely determined by the absolute elevation ψ and azimuth ϕ of the line-of-sight vector in terms of a fixed ground-based frame of reference.

that neglecting obstruction and higher-order effects, a SAR image is the coherent projection of the 3-D target complex reflectivity density into the slant plane, modified by waveform and processing artifacts. In particular, if the coherent projection of complex reflectivity density into the slant plane is denoted by $g_{\psi,\phi}(u,v)$, then the SAR image constructed according to above formulation (again neglecting obstruction and higher-order effects) can be shown to be a sampled version of

$$\hat{g}_{\psi,\phi}(u,v) = g_{\psi,\phi}(u,v) * \operatorname{sinc}\left(\frac{2\pi u}{\delta_r}\right) \operatorname{sinc}\left(\frac{2\pi v}{\delta_{xr}}\right), \tag{2.20}$$

where δ_r and δ_{xr} are defined in (2.13) and (2.19), respectively. If the SAR image is formed by sampling $\hat{g}_{\psi,\phi}(u,v)$ at a spacing finer than $\delta_r \times \delta_{xr}$ (*e.g.*, by zero-padding the sampled $G(f_x, f_y)$ data), so that the pixel spacing is finer than the resolution, then the image is said to be *oversampled*. The effect of oversampling is to provide a more smoothly interpolated representation of the projected reflectivity density.

One implication of (2.20) is that the locations of reflector primitives in SAR images are determined essentially by their projections into the slant plane. (To be precise, the locations of the responses of primitives in a SAR image are determined by the projections of their apparent reflection points into the slant plane. For point-like or flat reflector primitives such as corners, edges, flat plates, and dihedral- and trihedral-type bodies, the reflection points will remain constant across most viewing angles at which the primitive produces a bright response, but for reflector primitives with curved surfaces, such as cylinders and spheres, the reflection points will wander over the surface of the primitive as viewing angle changes [62].) Another implication of (2.20) is that bright scatterers will produce pronounced sidelobes in nearby pixels. To counteract this effect, the Fourier-domain phase history is often multiplied by a *sidelobe-suppression window function* prior to its DFT. In particular, if $\hat{G}(f_x, f_y)$ is the Cartesian representation of

$\hat{G}(f_r, f_\theta)$, then the phase history is formed by sampling

$$G(f_x, f_y) \cdot K_{ss}(f_x, f_y), \tag{2.21}$$

where $K_{ss}(f_x, f_y)$ is a unit-energy sidelobe-suppression window function with the same region of support as $G(f_x, f_y)$. (Typical choices for $K_{ss}(f_x, f_y)$ include Kaiser or Taylor windows [15].) This has the effect of modifying (2.20) so that

$$\hat{g}_{\psi,\phi}(u, v) = g_{\psi,\phi}(u, v) * \operatorname{sinc}\left(\frac{2\pi u}{\delta_r}\right) \operatorname{sinc}\left(\frac{2\pi v}{\delta_{xr}}\right) * k_{ss}(u, v), \tag{2.22}$$

where $k_{ss}(u, v)$ is the inverse Fourier transform of $K_{ss}(f_x, f_y)$. This convolution reduces the amplitude of the sidelobes but also reduces the effective resolution of the SAR image.

■ 2.3 Propagation Phenomenology

Clearly, in order to extract a target model from multiple SAR images, we require some phenomenological model relating the target to its appearance in images obtained at different aspect angles. The choice of any phenomenological model involves a tradeoff between accuracy and simplicity. For instance, given perfect knowledge of the transmitted field and all target characteristics, Maxwell's equations will describe the electric field at all points in space, and thus predict the radar image of the scene exactly. Unfortunately, the solution of Maxwell's equations for a complex target—let alone their inversion to yield a target model from scattered data—is generally intractable if not impossible.[10] It is therefore necessary to turn to a simpler model to predict the radar return from most targets [21, 61, 62, 92] and to facilitate the phenomenological model manipulation required for the inverse problem.

We will be concerned with systems operating at frequencies near 10 GHz, in the X band. The wavelength of a 10-GHz wave is 3 cm. If all target features are somewhat larger than this wavelength, then the interaction between the target and the electric field is said to occur in the *high-frequency* regime, and a high-frequency scattering model [61] may be utilized to predict the radar return. (High-frequency models depend on the approximation that all objects interacting with the electric field are substantially larger than the wavelength. So-called *low-frequency* and *resonant* models predict the interaction of electric fields with bodies that are substantially smaller than or approximately the same size as the wavelength, respectively [92].) We now describe three high-frequency models that are especially relevant to our efforts.

■ 2.3.1 Geometrical Optics

The simplest high-frequency model for RCS prediction is known as the Geometrical Optics (GO) model [62, 92]. GO was developed centuries ago and is based on Fermat's

[10] In fact, without regularizing assumptions of some kind, the inversion of Maxwell's equations to yield a target model from backscattered data is an ill-posed problem [59].

Principle (*i.e.*, corpuscles of light follow a path of minimum length) and Snell's Law (*i.e.*, the angle of reflection from a surface equals the angle of incidence). GO predicts that the backscattered RCS at any viewing angle will depend only on the geometry of the bodies illuminated by the radio waves and on their radii of curvature at the points of illumination. GO is useful primarily for predicting backscattering for *specular* reflections, that is, when reflected rays are parallel to the line of sight. GO is limited in several important respects: it fails to predict backscattering for non-specular reflections; it is applicable only when the radii of curvature at the points of reflection are finite and much larger than the wavelength; it is also inaccurate when the point of reflection lies near a surface edge or boundary. These failings limit the ability of GO to predict the scattering from a complex target, although, as we will see, it is successful in predicting the main components of the responses from primitives with a significant specular return, such as trihedrals and tophats.

■ 2.3.2 Physical Optics

A more sophisticated model for scattering and RCS prediction is known as Physical Optics (PO) [62, 92]. In the 1930s, Stratton and Chu noted that in the absence of sources, the electrical field at any point in space can be expressed as an integral over a closed surface containing the scattering volume. This observation, in addition to several simplifying assumptions [62], admits approximate closed-form backscattering solutions for some idealized primitives, such as spheres, cylinders, cones, and flat plates.[11] These approximate solutions do not require specular incidence and thus represent a significant improvement over GO.

PO provides an accurate prediction of the scattering from a much wider variety of targets and scenes than does GO, but still has failings: its accuracy deteriorates when incidence is very non-specular, and it cannot describe the contributions of edges and corners to the backscattered field. However, PO does provide a framework sufficient for the purposes of describing targets as a collection of reflector primitives of the type we are considering here—namely, finite 3-D bodies such as trihedrals, tophats, dihedrals and cylinders.

Consider the primitives depicted in Figure 2.10: a rectangular plate, a circular plate, a 90° rectangular-plate dihedral, a right circular cylinder, a 90° square-plate trihedral, and a circular-base tophat. Together, GO and PO can describe the primary response mechanisms of each of these primitives.[12] Each of these primitives has an intrinsic frame

[11] These assumptions also simplify the Stratton-Chu integral to the point where numerical integration and thus computer simulation becomes feasible; the scattering-prediction software package XPatch is largely based on PO [41].

[12] Here we are distinguishing "primary" response mechanisms, in which each of the primitive's components play a role in backscattering energy, from "secondary" response mechanisms, in which only a subset of the primitive's components have a role in backscattering. For instance, a trihedral's primary response mechanism is a triple-bounce return, in which an incoming ray is reflected off all three trihedral plates in turn before traveling back toward the antenna; its secondary mechanisms are the single- and double-bounce returns that arise when the trihedral is oriented at oblique angles and energy is

response mechanism	(ampl) · (pol) · (shaping function)
rectangular plate	$\left(\frac{k_0^2 a^2 b^2}{\pi}\right) \cdot \lvert\gamma_{tr}^{\mathrm{o}}\rvert^2 \cdot (\cos\psi' \cos\phi' \operatorname{sinc}(k_0 b \sin\psi') \operatorname{sinc}(k_0 a \cos\psi' \sin\phi'))^2$
circular plate	$(k_0^2 \pi r^4) \cdot \lvert\gamma_{tr}^{\mathrm{o}}\rvert^2 \cdot \left(\cos\psi' \cos\phi' \frac{J_1(2k_0 a(1-\cos^2\psi' \cos^2\phi')^{1/2})}{k_0 a(1-\cos^2\psi' \cos^2\phi')^{1/2}}\right)^2$
dihedral double-bounce	$\left(\frac{2k_0^2 a^2 b^2}{\pi}\right) \cdot \lvert\gamma_{tr}^{\mathrm{e}}\rvert^2 \cdot \left(\sqrt{2}\cos\psi' \sin\left(\frac{\pi}{4} - \lvert\phi' - \frac{\pi}{4}\rvert\right) \operatorname{sinc}(k_0 b \sin\psi')\, W_{[0,\frac{\pi}{2}]}(\phi')\right)^2$
trihedral triple-bounce	$\left(\frac{3k_0^2 a^4}{\pi}\right) \cdot \lvert\gamma_{tr}^{\mathrm{o}}\rvert^2 \cdot \begin{cases} \frac{1}{3}\left(\frac{4 l_i m_i}{n_i}\right)^2 W_{[0,\frac{\pi}{2}]}(\phi') W_{[0,\frac{\pi}{2}]}(\psi'), & m_i \leq \frac{n_i}{2}, \\ \frac{1}{3}\left(l_i\left(4 - \frac{n_i}{m_i}\right)\right)^2 W_{[0,\frac{\pi}{2}]}(\phi') W_{[0,\frac{\pi}{2}]}(\psi'), & m_i > \frac{n_i}{2}, \end{cases}$ where $(l_i, m_i, n_i) = \operatorname{sort}_{\min\to\max}(\sin\psi',\ \cos\psi'\sin\phi',\ \cos\psi'\cos\phi')$
cylinder shaft	$(k_0 r h^2) \cdot \lvert\gamma_{tr}^{\mathrm{o}}\rvert^2 \cdot (\cos\psi' \operatorname{sinc}(k_0 h \sin\psi'))^2$
tophat double-bounce	$(2k_0 r h^2) \cdot \lvert\gamma_{tr}^{\mathrm{e}}\rvert^2 \cdot \left(\sqrt{2}\sin\left(\frac{\pi}{4} - \lvert\psi' - \frac{\pi}{4}\rvert\right)\right)^2$

Table 2.1. Physical optics/geometrical optics RCS approximations for basic scattering mechanisms.

of reference, depicted by the sets of axes in Figure 2.10, whose rotation with respect to an absolute frame of reference, together with the absolute viewing angle, defines a relative viewing elevation ψ' and azimuth ϕ', also depicted in Figure 2.10. PO and GO predict the primary responses from these primitives as functions of relative viewing angle, the relevant primitive dimensions indicated in Figure 2.10, and the wavenumber k_0 of the illuminating radiation [62, 92], where

$$k_0 = \frac{2\pi c}{f_0} \tag{2.23}$$

by definition for radiation of frequency f_0. In particular, GO suffices to predict the primary trihedral and tophat responses, which are specular; PO is required to predict the other four responses. The GO and PO predictions for all six primitives are given in Table 2.1. Each response in the table is expressed as a product of three terms: a size-dependent amplitude coefficient, a polarimetry-dependent attenuation factor $|\gamma_{tr}^{\cdot}|^2$ between 0 and 1 that depends on the antenna transmit-receive polarization as described shortly, and a view-dependent shaping function that attains a maximum value of one. (Thus the maximum RCS obtained by any of the primitives in Figure 2.10 is given by the amplitude coefficients listed in Table 2.1.) The $W_{[\cdot,\cdot]}(\cdot)$ term appearing in several entries of Table 2.1 is the window function defined by (2.7) and the $J_1(\cdot)$ term appearing in the circular plate response is the Bessel function of order one.

Several features of the RCS models in Table 2.1 bear highlighting. First, the strength of the backscattered field (proportional to the square root of the RCS response) varies with the projection of the primitive into the plane orthogonal to the

backscattered by interaction with only one or two of its three plates.

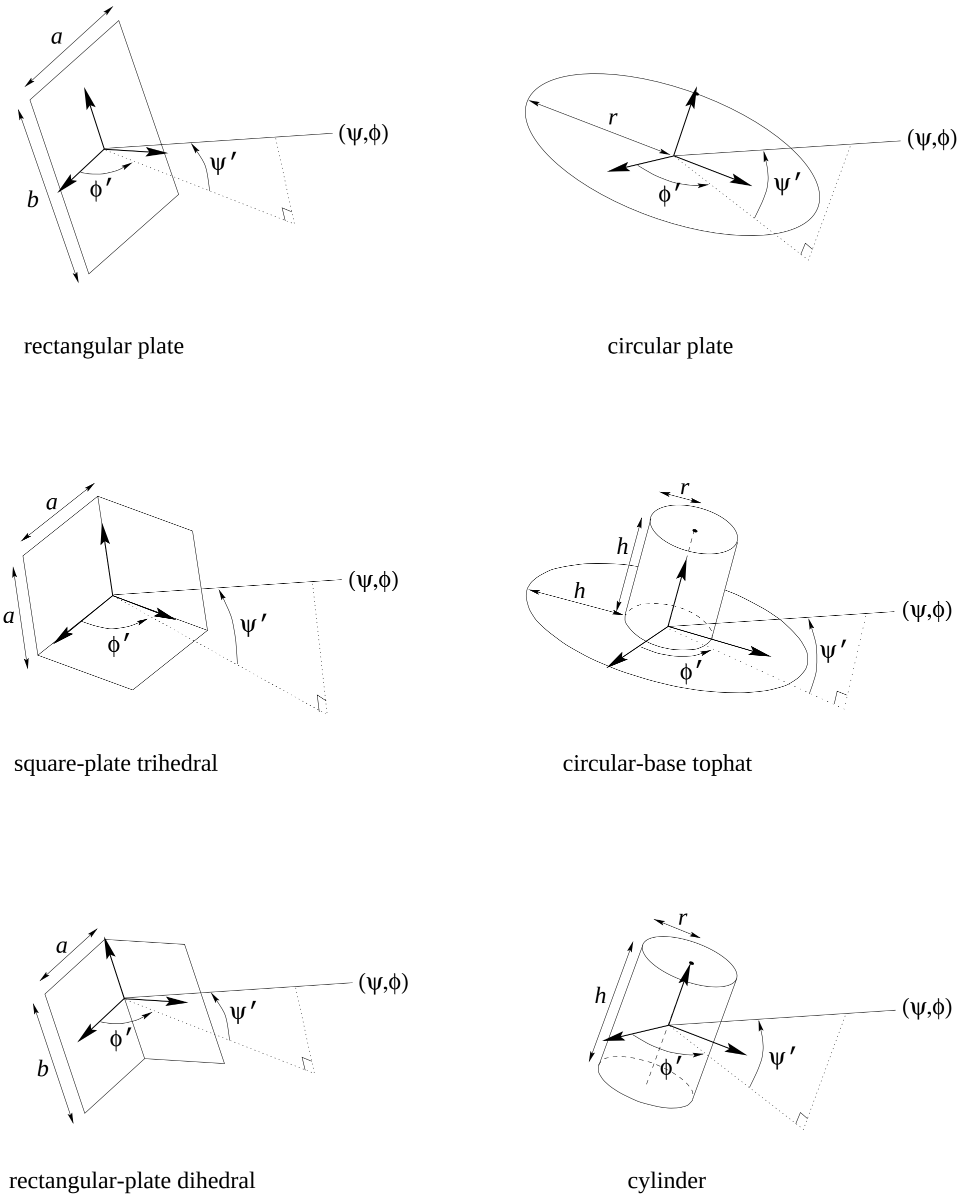


Figure 2.10. Basic primitives. Here (ψ, ϕ) is the absolute viewing elevation and azimuth, defined in terms of a fixed ground-based frame of reference as in Figure 2.9. This absolute viewing angle, together with the absolute orientation of a primitive, defines a relative viewing elevation ψ' and azimuth ϕ' as depicted above. The indicated dimensions and relative viewing angle of any primitive determine its RCS.

	HH	HV	VH	VV
$\gamma^{\mathrm{o}}_{\bullet}$	1	0	0	1
$\gamma^{\mathrm{e}}_{\bullet}$	$\cos 2\nu'$	$-\sin 2\nu'$	$\sin 2\nu'$	$-\cos 2\nu'$

Table 2.2. Polarization attenuation factors for basic scattering mechanisms.

line-of-sight vector. Second, while the responses of the rectangular and circular plates, dihedral, and cylinder fall off relatively quickly from their maximum values, the tophat and trihedral responses have much more angular breadth. This is because the unique geometries of the trihedral and tophat result in specular reflections over a wide range of viewing angles whereas the other primitives have a single specular plane or direction. Third, the sinc(•) and $J_1(\bullet)$ terms in the plate, dihedral, and cylinder responses are characteristic of PO predictions; note that in each case the off-broadside falloff is inversely proportional to the physical extent of the scatterer, so that larger primitives produce narrower returns.

As indicated above, the $|\gamma^{\bullet}_{tr}|^2$ terms in Table 2.1 capture the dependence of RCS on the polarizations of the transmitting and receiving antennas, as indicated by the tr subscripts. Physical optics predicts that the RCS will depend on whether an incoming ray is reflected an even or odd number of times—hence the two different polarization terms γ^{e}_{tr} and γ^{o}_{tr}. For even-bounce mechanisms, the relative rotation of the primitive also affects the RCS.[13] Denoting this rotation by ν', physical optics predicts the polarization attenuation factors given in Table 2.2 [92, 104]. This table indicates that while odd-bounce primitives reflect radiation without changing its polarization, even-bounce primitives rotate the polarization in the process of reflection; the pairs $(\gamma^{\mathrm{e}}_{HH}, \gamma^{\mathrm{e}}_{HV})$ and $(\gamma^{\mathrm{e}}_{VH}, \gamma^{\mathrm{e}}_{VV})$ indicate the components of the rotated polarization along orthogonal axes for horizontal and vertical transmit polarizations, respectively. Note that this rotation does *not* affect overall field strength. Note also that Table 2.2 implies that even- and odd-bounce mechanisms have distinctive polarimetric signatures that could be used to help distinguish between classes of primitives in cases when polarimetric radar measurements are available.

■ 2.3.3 The Geometrical Theory of Diffraction

In scenes with multiple large reflector primitives such as trihedrals, tophats, dihedrals, and cylinders, PO is generally sufficient to describe most of the observed backscattered energy. In situations where scattering contributions from edges or corners form a significant portion of the reflected energy, however, PO must be augmented if greater accuracy is desired. Although PO's predictions are adequate for most of the purposes of this thesis, we present one augmented model in the interest of thoroughness and as a prerequisite to the discussion of Section 2.4. This model is known as the Geomet-

[13]The rotation corresponds to Huynen's "tilt angle" [46] and is defined for a dihedral as the angle between the dihedral seam and a horizontal axis, and for a tophat as the angle between the tangential extension of the cylinder-base seam orthogonal to the specular direction.

reflector type	$\alpha(\cdot)$
flat plate at broadside	1
dihedral double-bounce at 0° depression	1
trihedral triple-bounce	1
cylinder at 0° depression	1/2
singly curved surface	1/2
tophat double-bounce	1/2
sphere	0
straight edge at broadside	0
point scatterer	0
doubly curved surface	0
curved edge at broadside	$-1/2$
corner	-1
tip	-1

Table 2.3. Reflectors and their GTD frequency-dependence parameters.

rical Theory of Diffraction (GTD) and was developed by Joseph Keller in the early 1960s [54]. GTD attempts to address the failing of PO to model edge and corner scattering contributions by extending the GO concept from two-dimensional surfaces to one-dimensional edges and point regions such as corners, tips, and vertices, predicting that fields incident upon these features are scattered along a "Keller cone" of directions. GTD succeeds in describing the frequency dependence of edge and corner contributions exactly but is silent regarding the angular variation of these responses. In particular, GTD predicts that the backscattered RCS observed from a feature illuminated at frequency f is proportional to $f^{2\alpha}$, where α is a feature-characteristic constant exponent that is an integer multiple of 1/2. Table 2.3 lists the GTD predictions of frequency dependence for various features [54, 85]. (Note the correspondence between the GTD and PO frequency-dependence predictions for the six primitives also listed in Table 2.1, remembering that frequency and wavenumber are inversely proportional.)

As indicated above, GTD does not provide descriptions of the angular variation in scattering observed from edges, corners, and related features. (There are extensions to GTD that attempt to predict the spatial dependence of edge and vertex scattering, but these are generally very complicated and often inaccurate outside of a narrow aspect region [61, 62].) As such, GTD's augmentation to PO is helpful primarily in scenarios where radar data are collected at a single aspect (as in a range profile) or over a range of aspects sufficiently narrow that any variation in amplitude is not pronounced (as in a narrow-aperture SAR image). This suggests that GTD's augmentations are of questionable value in our problem, which fundamentally requires a scattering model that accurately describes the variation in amplitude over widely separated aspect angles. GTD offers considerable value to single-aspect or narrow-aperture applications, however, as we describe in Section 2.4.

2.3.4 Phenomenology for Targets Comprised of Reflector Primitives

Together, the high-frequency scattering models of GO, PO, and GTD provide very accurate predictions of the observed backscattered responses from targets comprised of discrete reflector primitives (which might include not only components such as trihedrals and tophats, but also edges, corners, and tips). In particular, given a target comprised of M such discrete primitives, the backscattered response at frequency f, azimuth ϕ, and depression ψ can be expressed as

$$\mathcal{E}(f, \phi, \psi) = \sum_{m=1}^{M} A_m \left(j \frac{f}{f'} \right)^{\alpha(t_m)} e^{-j4\pi f r_m / c} \, s_{t_m}(\phi, \psi, \boldsymbol{\xi}_m), \tag{2.24}$$

where each one of the M reflector primitives is characterized by a set of parameters $\{A_m, t_m, r_m, \boldsymbol{\xi}_m\}$. The quantity r_m represents the down-range location of a given scatterer, A_m expresses the (possibly polarimetric) complex amplitude of the scatterer, and t_m is a discrete index identifying the fundamental type of the primitive (*e.g.*, flat plate or dihedral). The frequency dependence is captured by $\alpha(t_m)$, a mapping from discrete primitive type to the integer multiple of 1/2 characteristic of feature type t_m as indicated in Table 2.3. (Note that $\alpha(t_m)$ is a non-invertible mapping.) Information about reflector pose, shape, and size is encompassed in the abstract vector parameter $\boldsymbol{\xi}_m$ (the components of which might be specific to each primitive type). The quantity f' in (2.24) serves as a normalization term. The function $s_{t_m}(\phi, \psi, \boldsymbol{\xi}_m)$ describes the aspect dependence of the backscattering from a primitive of type t_m, and is provided by GO and PO for some primitive types (such as those in Table 2.1) but is generally unavailable for other primitive types (such as edges, corners, and tips).

The formulation of (2.24) is extremely useful and forms a basic model that enables the description of backscattering from a wide class of targets. Although our discussion has focused on a traditional set of primitives (augmented by features such as edges, corners, and tips), nothing in (2.24) requires such a close marriage to traditional components. Any feature that admits modeling by appropriate choices of $\alpha(t_m)$ and $s_{t_m}(\phi, \psi, \boldsymbol{\xi}_m)$ could be included as a reflector primitive. For instance, the class of primitives could be broadened to include responses from cavity structures or multiple-primitive reflections, as long as these phenomena lent themselves to a proper selection of $\alpha(t_m)$ and $s_{t_m}(\phi, \psi, \boldsymbol{\xi}_m)$. In practice, much of the research into target modeling and feature extraction has concentrated on the more traditional set of primitives described above. We now turn our attention to some of these efforts.

2.4 Previous Work

The last decade has seen a variety of attempts to model targets encountered in radar range profiles and SAR images. Much of this work has been done in the context of feature extraction for ATR; the reduction of a high-dimensional set of backscattered radar data to a handful of feature parameters by inversion of a model similar to that of (2.24) is attractive to ATR because it facilitates simple and efficient comparison of collected data and entries from the target database. In particular, if the target database is maintained as a set of feature vectors comprising parametric descriptions of targets, and if similar parametric target descriptions can be extracted from collected imagery or range profiles on the fly, then targets can be classified by comparison to database entries in the feature domain. This is central to the concept of model-based ATR.

In the following sections we describe recent efforts to extract parametric target descriptions, or feature vectors, from single-look radar data (*i.e.*, range profiles and SAR images). The focus of these efforts is conceptually similar to our problem: in both cases a scattering model must be specified and implicitly or explicitly inverted to yield a description of the target. Of course, there are also numerous differences in the two problems. One such difference is the requirement that our scattering model accurately describe backscattering aspect variation, a constraint largely absent in single-look problems. Another important difference is the difficult correspondence problem underlying the extraction of a target model from multiple viewing angles, as will be discussed in Chapter 3. Nevertheless, the work described here provides important insights into a similar problem. In Section 2.4.1 we describe efforts to extract target models from 1-D data, and discuss similar efforts to extract models from 2-D data in Section 2.4.2.

■ 2.4.1 Model Extraction from a Single Range Profile

One simple and obvious way to extract a model from a range profile is simply to apply an amplitude-based peak extractor or a matched filter to detect significant responses [23]. This method, while conceptually and computationally simple, has significant drawbacks. For instance, (2.13) states that the achievable Fourier range resolution is equal to $c/\eta_c T$; in practice it is often desirable to pinpoint the scatterer location to a finer scale. Additionally, a simple amplitude-based peak extractor would not provide type information.

One way of accomplishing this is to employ a parametric model for scattering behavior as in (2.24), in which scatterer locations and other information are extracted directly from radar data based on the assumed parametric forms of scatterers or primitives. In the literature this is often referred to as *scattering center extraction*. Use of a parametric model results in a *super-resolution* algorithm in the sense that its estimate of the location of a scatterer is not limited by the resolution of the range profile data.

Prony-Based Extraction

Some of the earliest relevant research into 1-D scattering center extraction [16, 45] was performed utilizing Prony's method [88]. Originally proposed in 1795, Prony's method is a procedure for finding a weighted sum of damped exponentials to fit a sequence of data. Prony's method has been widely studied and extended to a broad range of applications, and many good general references on the method exist [42, 68]. In the current context, the sequence of data is a sequence of stepped frequency measurements (*i.e.*, the Fourier transform of the range profile). Prony's method is used to find a fit to this data which is of the form

$$\mathcal{E}[n] = \sum_{m=1}^{M} a_m p_m^{f_n}, \tag{2.25}$$

where f_n is the nth stepped frequency Prony's method is used to obtain estimates of complex pole locations p_m (corresponding to the damped exponential modes) and their complex amplitudes a_m given an assumed model order M. Prony's method requires solving two $M \times M$ systems of linear equations and finding the zeros of an Mth order polynomial [68].

Although (2.25) appears to bear little resemblance to a phenomenologically based model such as (2.24), an appropriate change of variables [16] allows (2.25) to be rewritten as

$$\mathcal{E}[n] = \sum_{m=1}^{M} A_m q_m^{f_n} e^{-j4\pi f_n r_m/c}, \tag{2.26}$$

where A_m, q_m, and r_m are all functions of p_m and a_m from (2.25). Examination shows that (2.26) differs from (2.24) in only one respect: (2.26) predicts a frequency dependence of the form $q_m^{f_n}$, whereas (2.24) predicts a $f_n^{\alpha(t_m)}$ dependence. In situations where the *relative bandwidth* of the stepped frequency measurements is small, *i.e.*, where the range of stepped frequencies is small compared to the center frequency, the model of (2.26) might provide a reasonable approximation to (2.24), and Prony's method could then be used to extract estimates of scatterer location (r_m) and amplitude (A_m). (Note that, unfortunately, (2.26) does not enable direct estimation of scatterer type information, *i.e.*, the GTD frequency dependence exponent.)

Several authors have used Prony's method and the above transformation to facilitate extraction of target models from 1-D scattering data. In particular, Hurst and Mittra report on this approach in [45], assuming that the model order M is known *a priori*. They achieve success in locating scatterers to a precision unattainable by Fourier means. Additionally, they show that Prony's method provides a greater dynamic range in estimating the amplitudes of scattering centers than that provided by an amplitude-based peak extractor.[14] The work of [45] is extended by Carrière and Moses in [16], which presents a total-least-squares Prony (TLS-Prony) method [89] coupled with a singular-value decomposition to enable estimation of the model order. Carrière and Moses test their algorithm on radar data of aircraft, corrupted with additive white Gaussian noise, and show that they are able to extract scattering centers corresponding to physical target attributes such as tail fins, the rounded aircraft nose, and the leading edges of wings.

[14] The limitation in the dynamic range of an amplitude-based peak-extraction approach stems from the fact that the contributions of small-amplitude responses are often lost in the large sidelobes of the brighter responses.

GTD-Based Extraction

While the efforts described in [45] and [16] are both successful in estimating range locations and amplitudes of scatterers, they are deficient in several important respects. First of all, neither algorithm provides a means for estimating scatterer type; thus physical features of scatterers other than location and amplitude remain unknown.[15] Second, it is well known that Prony's method is very sensitive to noise [68], which is acknowledged in [45] and [16] although neither paper undertakes an analysis of performance under varying noise characteristics. Finally, as discussed above, Prony's method suffers from the inaccuracies inherent in its scattering model, which become more severe in systems with a large relative bandwidth. These deficiencies are the motivation for the work of [85], in which Potter *et al.* present an approximate maximum likelihood (ML) algorithm based on the model of (2.24) that builds on the TLS-Prony approach of [16] to arrive at estimates of $\{A_m, \alpha(t_m), r_m\}_{m=1}^M$. The algorithm is shown to produce parametric target descriptions with small squared error and accurate $\alpha(t_m)$ estimates given sufficiently large bandwidth. For small-bandwidth situations, estimation of $\alpha(t_m)$ is unreliable and it is not clear that estimation of r_m and A_m is any more reliable than in the Prony-based algorithm of [16]. The work of [85] demonstrates that although target models can be generated by inverting a GTD phenomenological model, inversion of a simpler Prony model might be preferable in narrow-bandwidth situations.

Dictionary-Based Extraction

In [70], McClure and Carin describe a quite general approach to scattering center extraction from 1-D data. Like [16,45,85], they model radar data as scaled superpositions of individual scatterer responses; unlike these previous efforts, however, their assumed scatterer responses are not tied to a single phenomenological model but are arbitrary basis functions chosen from a dictionary specified in advance. The dictionaries used in [70] are constructed to contain not only GTD-predicted responses (denoted in [70] as "wavefront" responses due to their relative localization in time) but also "resonant" responses (*i.e.*, responses localized in frequency that often arise from large-scale target-structure) and "chirp" responses (*i.e.*, responses localized in neither time nor frequency, often arising from cavities or dispersive materials on the target). Conceptually, their approach is general enough to allow consideration of any model that can be expressed in the form

$$\mathcal{E}[n] = \sum_{m=1}^{M} A_m d_m[n], \qquad d_m[n] \in \mathcal{D},\ m = 1, \dots, M \tag{2.27}$$

for an appropriately constructed (and reasonably sized) dictionary $\mathcal{D}$. Extraction of a scatterer model according to (2.27) requires selection of entries $d_m[n]$ from the dictionary $\mathcal{D}$ (presumably constructed so that any entry $d_m[n]$ has associated with it a set

[15] A later work by Potter *et al.* [85] uses estimates of $\alpha(t_m)$ derived from Prony p_m and a_m estimates, without specifying exactly how these were obtained.

of descriptive parameters, such as scatterer type and range) and a set of amplitudes A_m. McClure and Carin use matching pursuit techniques [67] to achieve this decomposition. Matching pursuit is a nonlinear technique developed by Mallat that is similar to projection pursuit [44], which projects elements from an overcomplete dictionary of basis elements onto a signal to obtain a compact representation of the signal. Matching pursuit is an iterative greedy algorithm which, at each iteration, augments its representation of the signal by choosing the dictionary entry with maximum projection onto part of the signal.

The selection of GTD-inspired wavefront responses for the dictionaries of [70] is straightforward: McClure and Carin generate basis elements by essentially sampling over the space of all feasible scatterer ranges r_m and frequency exponents $\alpha(t_m)$. The selection of appropriate resonance and chirp responses is more problematic, and it is argued in [70] that population of this part of the dictionary should be guided by *a priori* knowledge of target shape and structure, since an unconstrained selection of all feasible chirps and resonances would result in a prohibitively large dictionary.

McClure and Carin present results corresponding to two scenarios: one in which the target produces both wavefront and resonant responses, and one in which the dominant responses are chirps. Their results indicate that when the resonant and chirp dictionary elements are chosen based on exact knowledge of the resonant modes and dispersive character of the target, matching pursuit does a good job of selecting dictionary elements to provide a concise and accurate description of the target. However, they do not present results in the case where target resonance and dispersion are unknown or not precisely known in advance. McClure and Carin acknowledge that a lack of prior knowledge of target structure might impose restrictions on the applicability of the algorithm, although they claim that in applications in which a small number of possible targets might be encountered, the resonant and chirp dictionary elements can be reliably calculated in advance; Mallat also notes in [67] that in general some prior knowledge of signal structure is required in order to construct an effective matching pursuit dictionary. Depending on the model generation scenario, this may or may not be a significant limitation: sufficient prior knowledge might be provided by knowledge of the general class of targets being imaged, or by estimation of resonant modes or chirp characteristics directly from observed data before forming a dictionary. In any case, the approach proposed in [70] is general enough to be applicable in many situations.

2.4.2 Model Extraction from a Single SAR Image

All of the previous work discussed so far concerns scattering center extraction from 1-D radar data, that is, from range profiles or stepped frequency measurements. We now turn our attention to recent efforts to extract scattering centers from 2-D radar data in the form of SAR images or phase histories. As in the 1-D case, one conceptually simple approach is simply to apply an amplitude-based peak extractor to a SAR image to obtain estimates of scatterer locations and amplitudes; also as in the 1-D case, this approach has the drawbacks of limited resolution and no direct way to obtain type information (except in the case of polarimetric imagery). The work described here focuses on super-resolution techniques based on 2-D analogs to the Prony and GTD models.

Prony-Based Extraction

The extension of the Prony model of (2.25) to two dimensions is

$$\mathcal{E}[m,n] = \sum_{i=1}^{I} \sum_{j=1}^{J_i} a_{i,j} p_{x_i}^{f_m} p_{y_{i,j}}^{n}, \tag{2.28}$$

where p_{x_i} is the ith of I poles in the x component, $p_{y_{i,j}}$ is the i,jth pole in the y component (*i.e.*, the jth of J_k poles corresponding to the ith pole in the x component), and $a_{i,j}$ is the complex amplitude of the i,jth pole. (The quantity f_m is the mth stepped frequency, and n is the aspect index.) Note that this model is separable in the m and n indices. Like the 1-D model, it suffers from a mismatch in predicted frequency dependence that becomes more marked as relative bandwidth increases. Note that (2.28) models the spatial variation of the scattering amplitude across the imaging aperture as an exponential; this is due to the fundamental structure of the Prony model and not to any physical consideration, although over a narrow aperture an exponential term may be sufficient to model many primitives' scattering variations.

In [93], Sacchini, Steedly, and Moses present two methods for estimating I, $\{J_i\}_{i=1}^{I}$, and the complete set of p_{x_i}, $p_{y_{i,j}}$, and $a_{i,j}$ from two-dimensional data. (One of these methods is less accurate but computationally simpler than the other.) They also present an example of the application of the more accurate method to the extraction of scattering center range and cross-range locations from a SAR phase history. Their methods are similar in spirit to the 1-D Prony method of [16], employing a 2-D TLS-Prony technique to estimate model order, locations, and amplitudes. Like the 1-D Prony-based algorithms of [45] and [16], this 2-D Prony-based algorithm is incapable of estimating scatterer type.[16] The authors demonstrate the utility of their algorithm in extracting scatterer location and amplitude from a synthetic image constructed using a fairly large aperture and relative bandwidth.

GTD-Based Extraction

As in the 1-D case, the 2-D Prony model is an inaccurate description of the true scattering behavior that becomes more flawed as relative bandwidth or aperture increases. This limitation, and the desire to estimate scatterer type, motivated Potter and Moses to develop a GTD-based algorithm for scattering center extraction from SAR phase

[16] Potter and Moses [86] describe how type estimates can be obtained from the Prony-based algorithm in a later work.

histories [86]. They utilize a 2-D GTD model of the form

$$\mathcal{E}(f,\theta) = \sum_{m=1}^{M} A_m e^{\beta_m \theta} \left(\frac{jf}{f_c}\right)^{\alpha(t_m)} \exp\left\{\frac{-j4\pi f}{c}(x_m \cos\theta + y_m \sin\theta)\right\}, \tag{2.29}$$

where θ captures the aspect variation along the synthetic aperture, $\alpha(t_m)$ is the parameter describing the GTD-predicted frequency dependence of the mth scattering center, x_m and y_m are respectively the slant-plane range and cross-range locations of scatterer m that together determine its distance, and the aspect-dependent backscattering term of (2.24) has been modeled as $A_m e^{\beta_m \theta}$, a simple model intended to capture any observed amplitude variation over narrow apertures. (As in the 1-D model, A_m is possibly polarimetric.)

Potter and Moses describe an approximate ML algorithm (which is more completely described by Chiang in [18]) to estimate the parameters A_m, x_m, y_m, $\alpha(t_m)$, and β_m from (2.29) given M. The algorithm is similar in character to the 1-D GTD-based scattering center extraction algorithm in [85] described in Section 2.4.1. Potter and Moses present experimental results demonstrating the performance of their algorithm; as in [85], estimation of $\alpha(t_m)$ suffers when relative bandwidth is small. As in the case of [85], this suggests that the simpler Prony-based algorithm of [93] might be more appropriate for feature extraction except in cases where a large bandwidth offers significant value-added.

Another contribution of [86] is the demonstration of an M-ary generalized likelihood-ratio test (GLRT) that can be used to estimate reflector type from polarimetric data, independent of the GTD model of (2.24). This GLRT relies on the distinctive polarimetric signatures of different primitives [104] and can be used to obtain a finer estimate of reflector primitive type than is given by $\alpha(t_m)$ alone, in the sense that it can distinguish between many reflector types that share the same frequency-dependence exponent. The GLRT requires compilation of the polarimetric amplitude signatures for each candidate reflector primitive type[17] at rotational angles sampled over a full 360° of rotation about the line-of-sight vector. Their GLRT also yields as a byproduct an estimate of the in-plane rotational orientation of the reflector with respect to the look vector. They present one example, in which the GLRT correctly classifies two cylindrical objects and a trihedral. This is an example of how information about reflector type can be obtained without relying on frequency-dependence information, which is basically unreliable for narrow-bandwidth images.

[17] Potter and Moses use eight primitive types in the GLRT of [86], including trihedral, dihedral, narrow diplane, dipole, cylinder, quarter-wave plate, left-handed helix, and right-handed helix.

Chapter 3

Precise Problem Formulation

IN this chapter we establish a precise formulation of the problem of target model generation. Previous chapters have presented the general context in which target model generation is important, and future chapters will describe in detail our approach to the estimation of target models. This chapter serves a bridge between these two components by casting the problem in concrete terms. In particular, this chapter includes a formalization of the problem statement, a presentation of our conceptual approach to target model generation, an enumeration and analysis of our assumptions, and the construction of a measurement model enabling the estimation of a target model from the observed SAR data.

The data set from which we must estimate a target model is a collection of K spotlight-mode SAR images [15,51]. Each of these images is characterized by a particular viewing geometry, as depicted in Figure 2.9. In particular, the viewing geometry is defined by the line-of-sight vector from the center of a narrow synthetic aperture to the center of the target region being imaged. This line-of-sight vector can in turn be characterized by an azimuth and elevation, each defined with respect to an absolute ground-based frame of reference, as depicted in Figure 2.9. We assume no restriction on the azimuth ϕ_k and elevation ψ_k defining the line-of-sight vector for any image k; we do assume, however, that each image has been formed at a squint angle of $90°$ (see Figure 2.9), although extension of our approach to allow arbitrary squint angles is straightforward. Finally, we assume that each image was formed over a linear aperture in level flight, so that any (ψ_k, ϕ_k) pair defines a unique slant plane associated with image k.

A block-diagram representation of our general approach to 3-D target model estimation is depicted in Figure 3.1. A target is observed through a set of K SAR images. Each of these images is processed to extract a set of features which are then fused in order to produce a 3-D target model. The framework depicted in Figure 3.1 is quite flexible. For instance, we could consider modeling the target as a spatially varying scattering medium and use a trivial feature extraction stage that simply passes on each complete SAR image to the final module, which would then bear the full burden of inverse scattering. However, as indicated in the previous chapters, we have in mind a much more constrained approach that restricts our description of the target in order to focus on variables that are most observable and of significant interest to model-based

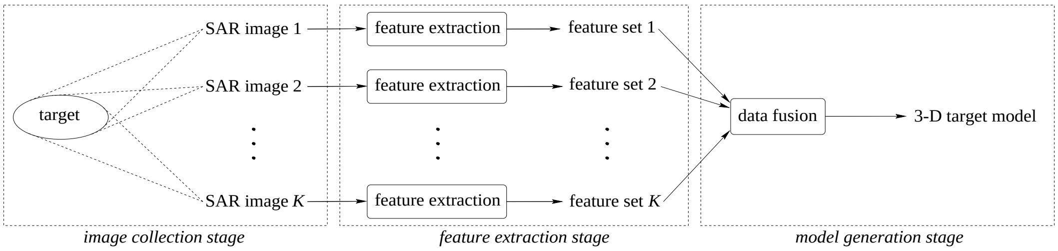


Figure 3.1. Target model generation block diagram.

ATR—namely, reflector-primitive-based parameterizations of targets.

Our principal focus is the design of the data fusion module in Figure 3.1. 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 former step involves judicious selection of parameters that are compact, descriptive, and observable; the latter step involves modeling both the SAR image collection process and the subsequent feature extraction stage that produces the observable data on which the fusion module will operate. These are the goals of this chapter. We describe the notation and basic target assumptions defining the problem in Section 3.1 and present our feature extraction procedure in Section 3.2. In Section 3.3 we construct a measurement model that relates the observable features and the 3-D target model parameters to be estimated.

■ 3.1 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 enable description of the scattering behavior of the primitive given any imaging geometry. In particular, a target model will be specified by the number of primitives N comprising the target and by N vectors of parameters, one associated with each component primitive. In general, we can express each vector $\boldsymbol{\theta}_i$ (for $i = 1, \ldots, N$) as

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_i^t \\ \theta_i^{\mathbf{x}} \\ \boldsymbol{\theta}_i^{\mathbf{d}} \end{bmatrix}, \tag{3.1}$$

where θ_i^t is an integer index designating the primitive as one of n_t canonical primitive types, $\theta_i^{\mathbf{x}}$ is a vector in $\mathbb{R}^3$ describing the location of the primitive, and $\boldsymbol{\theta}_i^{\mathbf{d}}$ is a generic vector parameter corresponding to a set of continuous- or discrete-valued descriptors that, along with θ_i^t and $\theta_i^{\mathbf{x}}$, specify the appearance of the primitive in an arbitrary SAR image. We will denote the log-RCS of 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.[1]

Most of our attention in this thesis is restricted to a small class of primitives including trihedrals, tophats, dihedrals, and cylinders (so that $n_t = 4$), depicted as the bottom four primitives in Figure 2.10; we assign type indices (*i.e.*, θ_i^t values) of 1 through 4 to these primitives, respectively. Location $\theta_i^{\mathbf{x}}$ of each primitive is defined to correspond to the origin of the primitive's local axes in Figure 2.10. A complete parameterization of any one of these primitives from the perspective of PO (see Section 2.3.2) would select $\boldsymbol{\theta}_i^{\mathbf{d}}$ to augment these parameters with all others required to describe its backscattering at an arbitrary viewing angle as predicted by Table 2.1—for instance, an Euler-angle pose [100] describing the primitive's orientation,[2] the size-dependent amplitude coefficient describing the primitive's maximum RCS, and any other dimensions of a primitive of type θ_i^t that are required to specify its shaping function exactly (*e.g.*, the dihedral and cylinder heights). Of course, other or more extensive parameterizations are possible. For instance, an augmented parameterization for characterization of nonideal primitives might include a vector of descriptive parameters (possibly specific to each primitive type) that enable modification to the canonical PO responses. A parameterization that modeled primitives as imperfect electrical conductors might include dielectric or absorption terms. In Chapters 7 and 8 we consider how the basic parameterization described here might be expanded to enable description of a wider variety of primitives.

Our basic parameterization is a slight compression of the "complete PO" parameterization that excludes primitive dimension information not relevant to determination of the primitive's maximum RCS or its slant-plane location. In particular, our basic parameterization augments θ_i^t and $\theta_i^{\mathbf{x}}$ with a $\boldsymbol{\theta}_i^{\mathbf{d}}$ consisting of at most three parameters: an overall *base amplitude* θ_i^a related to the overall size of the primitive and corresponding to the log of the amplitude coefficient in Table 2.1, an Euler-angle pose $\theta_i^{\mathbf{p}}$ indicating the orientation of the scatterer, and a radius of curvature θ_i^r for radially symmetric primitives including tophats and cylinders. The complete vector $\boldsymbol{\theta}_i$ in this parameterization is thus given by

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_i^t \\ \theta_i^{\mathbf{x}} \\ \theta_i^a \\ \theta_i^{\mathbf{p}} \\ \theta_i^r \end{bmatrix}. \tag{3.2}$$

The vector $\boldsymbol{\theta}_i$ provides a concise yet accurate description of a primitive's appearance in an arbitrary SAR image according to the physical and imaging models of Chapter 2:

[1] If polarimetric measurements are used, we take the scalar $A(\boldsymbol{\theta}_i, \psi, \phi)$ to be the log of the magnitude of the polarimetric vector.

[2] Three distinct Euler angles, corresponding essentially to elevation, azimuth, and rotation, are required to describe the orientation of primitives without radial symmetry, such as trihedrals and dihedrals. Two Euler angles suffice to describe the orientation of primitives with rotational symmetry, such as cylinders and circular-base tophats.

notation	description	space drawn from
θ_i^t	basic scattering type	$\{1,2,3,4\}$
$\theta_i^{\mathbf{x}}$	location	$\mathbb{R}^3$
θ_i^a	maximum RCS from any viewing angle	$\mathbb{R}$
$\theta_i^{\mathbf{P}}$	Euler-angle pose	$\begin{cases}[0,\pi]\times[0,2\pi)\times[0,2\pi), & \theta_i^t\in\{1,3\},\\ [0,\pi]\times[0,2\pi), & \theta_i^t\in\{2,4\}\end{cases}$
θ_i^r	radius of curvature	$\begin{cases}\emptyset, & \theta_i^t\in\{1,3\},\\ \mathbb{R}_+, & \theta_i^t\in\{2,4\}\end{cases}$

Table 3.1. Primitive parameterization.

location $\theta_i^{\mathbf{x}}$ and radius of curvature θ_i^r (for those primitives for which it is defined), along with viewing angle, determine the apparent location of the primitive in the SAR slant plane; discrete type θ_i^t identifies the primitive and selects a basic scattering response from Table 2.1; pose $\theta_i^{\mathbf{P}}$ 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 size of the primitive. The components of $\boldsymbol{\theta}_i$ in (3.2) are summarized in Table 3.1.

The physical and imaging models of Chapter 2 allow quantification of the previous statements. For instance, given the slant-plane-based SAR imaging model of Section 2.2, we can model the observed location of primitive i in image k as a measurement of $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$, the projection of the primitive's apparent primary-response specular-reflection point into the slant plane.[3] In particular,

$$\boldsymbol{\pi}_k(\boldsymbol{\theta}_i) = \begin{cases} \mathbf{H}_k\theta_i^{\mathbf{x}}, & \theta_i^t \in \{1,3\}, \\ \mathbf{H}_k\theta_i^{\mathbf{x}} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \theta_i^r \cos\psi'_{i,k}, & \theta_i^t \in \{2,4\}, \end{cases} \tag{3.3}$$

where $\psi'_{i,k}$ is the relative viewing elevation for primitive i at the center viewing angle of image k as pictured in Figure 2.10, and where $\mathbf{H}_k$ is the 2×3 ground-to-slant-plane transformation matrix for image k defined by the center viewing angle (see Figure 2.9):

$$\mathbf{H}_k = \begin{bmatrix} -\cos\psi_k\cos\phi_k & \cos\psi_k\sin\phi_k & -\sin\psi_k \\ -\sin\phi_k & -\cos\phi_k & 0 \end{bmatrix} \tag{3.4}$$

Similarly, the PO models of Section 2.3.2 allow us to specify an equation for the RCS of primitive i at the center viewing angle of image k:

$$A(\boldsymbol{\theta}_i, \psi_k, \phi_k) = \theta_i^a + S_{\theta_i^t}(\psi'_{i,k}, \phi'_{i,k}) + c_{\mathrm{pol}}, \tag{3.5}$$

[3] This model will be flawed, of course, at viewing angles from which secondary reflection mechanisms dominate, or from which the response is not in the specular mainlobe [62]. Although we could construct a model that takes these effects into account, doing so would significantly complicate the ensuing development. The implications of using a simpler model such as the one that follows are examined in Chapter 5.

where $\phi'_{i,k}$ is the relative viewing azimuth for primitive i in image k (see Figure 2.10) and θ_i^a, $S_{\theta_i^t}(\cdot)$, and c_{pol} have been constructed from the components of Table 2.1: θ_i^a is the log of the premultiplying amplitude coefficient for a primitive of type θ_i^t, $S_{\theta_i^t}(\cdot)$ is constructed from the log-shaping functions of the primary and secondary response mechanisms for a primitive of type θ_i^t as described shortly, and c_{pol} captures the polarimetry scaling factor and includes any effective gain achieved by using multiple polarimetric channels. For each primitive type, $S_{\theta_i^t}(\cdot)$ is scaled to give a maximum value is 0 dBsm, so that θ_i^a will correspond to the maximum single-polarization RCS of the primitive response.

As indicated above, the basic parameterization of (3.2) neglects the dependence of primitives' PO-predicted shaping functions in Table 2.1 on individual dimensions. This dependence is most pronounced for the dihedral and cylinder, which exhibit sinc-like elevation responses depending on b and h, respectively (see Figure 2.10). This dependence is also present, though less pronounced, for the trihedral and tophat: although their primary response mechanisms are specular reflections whose shaping functions are not dependent on individual primitive dimensions, their secondary response mechanisms incorporate dihedral and cylinder reflections, respectively (as well as plate responses). Additionally, primitive responses that comprise two or more 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 $S_{\theta_i^t}(\cdot)$ on nominally chosen dimensions for each primitive, even though real scatterers may have signatures that deviate from these $S_{\theta_i^t}(\cdot)$. 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 $\boldsymbol{\theta}_i^{\mathbf{d}}$ to include all relevant dimensions for each primitive type and construct $S_{\theta_i^t}(\cdot)$ with appropriate dependences on these dimensions. Our basic approach is the first of these alternatives. In particular, we construct our $S_{\theta_i^t}(\cdot)$ using nominal values for each primitive's relevant dimensions, chosen to correspond to moderately bright primitives. (The precise values chosen for use throughout most of this thesis are presented in Chapter 5, in the context of the experiments in which they are first used.) These nominal dimensions 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 function $S_{\theta_i^t}(\cdot)$ for each primitive type. Details of our construction of these scattering functions are given in Appendix A. Extensions to the second or third alternatives listed above are conceptually straightforward (although with a computational cost). In Chapter 5 we explore to what extent a mismatch between actual primitive dimensions and the nominally chosen dimensions selected to construct the $S_{\theta_i^t}(\cdot)$ affects performance.

Before proceeding, we introduce a vector $\boldsymbol{\theta}$ encompassing the complete parametric description of the target:

$$\boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \vdots \\ \boldsymbol{\theta}_N \end{bmatrix}. \tag{3.6}$$

(Note that $\boldsymbol{\theta}$ implicitly specifies the model order N.) Target model generation in our framework is thus estimation of the vector $\boldsymbol{\theta}$. We will model $\boldsymbol{\theta}$ (and thus N) as unknown parameters about which no information is available other than that provided by the SAR images.

■ 3.2 Extracted Features for Target Model Generation

As described previously and depicted in Figure 3.1, our framework entails selection of a feature extraction scheme to compress the full set of SAR imagery into a vector of descriptive parameters. These descriptive parameters must bear a clear relation to the components of $\boldsymbol{\theta}$ if model estimation is to be achieved. This implies that at the very least, the features extracted from each image should include some description of the locations and amplitudes of the dominant scatterers. In other words, the chosen features should include measurements of $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$ to provide information about $\theta_i^{\mathbf{x}}$ and θ_i^r as in (3.3), and measurements of $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ to provide information about θ_i^t, θ_i^a, and $\theta_i^{\mathbf{P}}$ as in (3.5). Additionally, if polarimetric data are available, the feature set should include some measurement of the polarimetric signature, since this clearly relates to θ_i^t as discussed in Section 2.3.2.

The full data set from which selected features must be extracted is the set of K spotlight-mode SAR images. We assume that each of these images is polarimetric so that at any pixel $[m, n]$ we have available a vector measurement of the form

$$\mathbf{g}[m,n] = \begin{bmatrix} g_{HH}[m,n] \\ \sqrt{2}g_{HV}[m,n] \\ g_{VV}[m,n] \end{bmatrix}, \tag{3.7}$$

where H and V are horizontal and linear polarizations as defined in Section 2.1.1 (see Figure 2.1).[4] 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 scene center) are known and can be related to the absolute ground-based frame of reference. Such information could be provided, for instance,

[4] Although a SAR imaging system employing two distinct linear antenna polarizations could in theory return four distinct polarimetric channels (HH, HV, VH, and VV), in practice most such systems return only three distinct channels as indicated above. This is an effect of polarimetric-calibration processing in which the inherent symmetry between HV and VH polarizations [104] is used to correct phase errors between the polarimetric channels [23]. The HV and VH responses are combined into a single channel (typically designated HV), leading to the cross-polarization channel gain factor of $\sqrt{2}$ in (3.7).

by geolocation or global positioning measurements taken as the images are collected,[5] coupled with accurate ranging and positioning of the target.

The 2-D extraction techniques of Section 2.4.2 provide potential feature extraction schemes: the Prony-based technique of [93] provides location and amplitude information, and the GTD-based technique of [86] provides location, amplitude, and type information (in wide-bandwidth images). For the sake of simplicity and clarity, we choose to utilize a simpler feature extraction technique than those provided by [86, 93]. In particular, we utilize an amplitude-based peak extractor that directly provides location and amplitude information for the brightest scatterers in an image; this information is augmented by type information obtained from the polarimetry measurements. This approach is conceptually straightforward and has the important practical advantage of potentially being easier to characterize statistically than the approaches of [86, 93]. (Statistical characterization of the feature extraction process is essential for the development of a measurement model, as described in Section 3.3.) One drawback to this approach is its inability to pinpoint scatterer location to the degree attainable by the techniques described in [86, 93]. This disadvantage is mitigated somewhat by the fact that we will likely have multiple location measurements of each primitive from many aspects with which to estimate a location in $\mathbb{R}^3$.

Our peak extractor compresses each SAR image into a handful of augmented detections corresponding to the locations, amplitudes, and polarimetric signature classifications of the brightest responses in the image. The input to the peak extractor is the set of polarimetric data $\mathbf{g}[m,n]$ in (3.7) for all pixels $[m,n]$. We can form a single magnitude image $\bar{g}[m,n]$ from the pixelwise magnitudes of these polarimetric vectors, *i.e.*,

$$\bar{g}[m,n] = \|\mathbf{g}[m,n]\| \tag{3.8}$$

for any pixel $[m,n]$. The peak extractor identifies peaks by selecting pixels from $\bar{g}[m,n]$ that are local maxima with amplitude greater than some specified threshold. This produces an arbitrary number of peaks or detections in any image k, denoted M_k; each peak $j = 1, \ldots, M_k$ in image k corresponds to a distinct pixel $[m_{k,j}, n_{k,j}]$. Our feature extraction technique describes each peak j in each image k with three measurements: a 2-D slant-plane range/cross-range subpixel location $\mathbf{x}_{k,j}$ that is a measurement of $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$ in (3.3), a scalar amplitude $a_{k,j}$ that is a dBsm measurement of $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ in (3.5), and a binary polarimetric-type-classification $t_{k,j}$ that is a measurement of the polarimetric signature of (3.7). These measurements are obtained as follows. The amplitude is simply taken as a dBsm measurement from $\bar{g}[m,n]$, *i.e.*,

$$a_{k,j} = 20 \log_{10} \bar{g}[m_{k,j}, n_{k,j}]. \tag{3.9}$$

[5]Existing SAR systems typically already make such measurements. Slight deviations from a linear constant-velocity flightpath are generally recorded so that their effects may be corrected in the image formation process [51].

Location $\mathbf{x}_{k,j}$ is a subpixel estimate of scatterer slant-plane location obtained using a simple parabolic interpolation procedure [87] based on $\bar{g}[m_{k,j}, n_{k,j}]$ and its four nearest neighbors:

$$\mathbf{x}_{k,j} = \mathbf{T} \begin{bmatrix} m_{k,j} - \frac{\bar{g}[m_{k,j}+1,n_{k,j}]-\bar{g}[m_{k,j}-1,n_{k,j}]}{2(\bar{g}[m_{k,j}+1,n_{k,j}]+\bar{g}[m_{k,j}-1,n_{k,j}]-2\bar{g}[m_{k,j},n_{k,j}])} \\ n_{k,j} - \frac{\bar{g}[m_{k,j},n_{k,j}+1]-\bar{g}[m_{k,j},n_{k,j}-1]}{2(\bar{g}[m_{k,j},n_{k,j}+1]+\bar{g}[m_{k,j},n_{k,j}-1]-2\bar{g}[m_{k,j},n_{k,j}])} \end{bmatrix}, \tag{3.10}$$

where $\mathbf{T}$ is a 2×2 matrix transforming pixel indices into range/cross-range coordinates. (This quadratic interpolation is motivated by the observation that the mainlobe of the convoluting kernel in (2.22) can be modeled as quadratic near its peak for any standard SAR sidelobe-suppression window function [15].) The type measurement $t_{k,j}$ is a binary classification of $\mathbf{g}[m_{k,j}, n_{k,j}]$ as either an odd-bounce or even-bounce signature as predicted by PO (see Table 2.2), obtained via a polarimetric-signature GLRT classifier similar to those of [32,86].[6] Specifically, the classifier's GLRT signature hypotheses are

$$\begin{aligned} H_{\text{odd}}: \qquad \mathbf{g} &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \\ H_{\text{even}}(\nu'_n): \qquad \mathbf{g} &= \begin{bmatrix} \cos 2\nu'_n \\ \sqrt{2}\sin 2\nu'_n \\ -\cos 2\nu'_n \end{bmatrix}. \end{aligned} \tag{3.11}$$

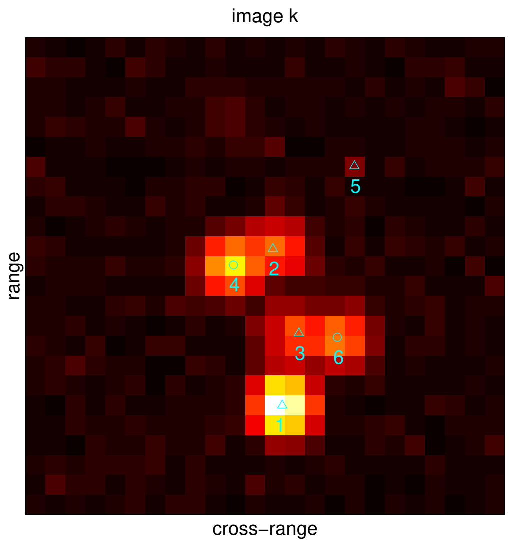
(We generate even-bounce hypotheses to span a $180°$ interval at $5°$ increments, so that $\nu'_n = (5n)°$ for $n = 0, 1, \ldots, 35$.) We designate an odd-bounce classification as $t_{k,j} = 1$ and an even-bounce classification as $t_{k,j} = 2$.

Figure 3.2 is an illustration of the features extracted from a typical SAR image chip—in this case, an XPatch-generated polarimetric image (presented here in log-magnitude form) of a SLICY-like target in the presence of background clutter. Note that the peak extractor identifies and describes the clearly salient features of the image, effectively compressing a high-dimensional data set (24×24 pixels of three polarimetric measurements each, for a total of 1728 measurements) into a handful of descriptive parameters.

For convenient reference, the three-parameter location/amplitude/type description of the jth peak of image k will be called a *report* and denoted by $\mathbf{Z}_{k,j}$, which can be expressed as

$$\mathbf{Z}_{k,j} = \begin{bmatrix} \mathbf{x}_{k,j} \\ a_{k,j} \\ t_{k,j} \end{bmatrix}. \tag{3.12}$$

[6] This GLRT provides information about the orientation of even-bounce scatterers (namely, the rotation angle ν' in Table 2.2) as a byproduct of its classification. Although this information could be used to supplement the features described here, we do not include it in the feature set.



j	$\mathbf{x}_{k,j}$ ([m, m]^T)	$a_{k,j}$ (dBsm)	$t_{k,j}$
1	$[-1.30,\ -0.17\ ]^T$	21.53	1
2	$[\ \ 0.27,\ -0.08\ ]^T$	16.03	1
3	$[-0.58,\ -0.34\ ]^T$	15.00	1
4	$[\ \ 0.11,\ \ \ 0.32\ ]^T$	18.79	2
5	$[\ \ 1.11,\ -0.90\ ]^T$	6.34	1
6	$[-0.62,\ -0.72\ ]^T$	16.10	2

Figure 3.2. Feature extraction from a SAR image chip.

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,

$$\mathbf{Z}_k = \begin{bmatrix} \mathbf{Z}_{k,1} \\ \vdots \\ \mathbf{Z}_{k,M_k} \end{bmatrix}, \tag{3.13}$$

and notation for all reports in all images,

$$\mathbf{Z} = \begin{bmatrix} \mathbf{Z}_{1} \\ \vdots \\ \mathbf{Z}_{K} \end{bmatrix}. \tag{3.14}$$

The vector $\mathbf{Z}$ thus represents the complete set of observations available to the data fusion module in Figure 3.1, *i.e.*, the complete set of extracted features from which $\boldsymbol{\theta}$ is to be estimated.

■ 3.3 Measurement Model

Feature extraction is a stochastic process: it takes place in the presence of scene clutter and background noise. It is subject to missed detections, false alarms, and measurement uncertainty. Any model relating $\mathbf{Z}$ to $\boldsymbol{\theta}$ must describe this uncertainty. In this section we turn our attention to the construction of a probabilistic measurement model that will serve as the basis for estimation of $\boldsymbol{\theta}$ from $\mathbf{Z}$.

The uncertainties in the extracted features fundamentally come at two levels of granularity, one coarse and one fine. The coarse-level uncertainty involves the identity

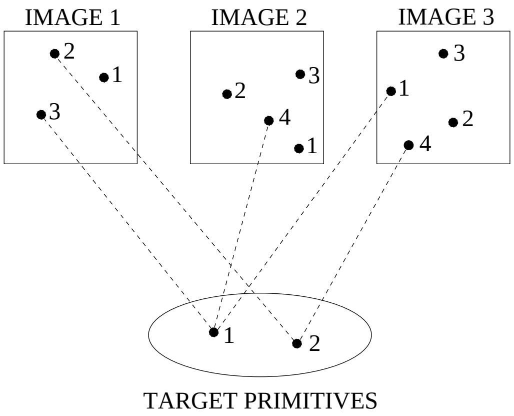


image 1	image 2	image 3
$M_1 = 3$	$M_2 = 4$	$M_3 = 4$
$\lambda_{1,1} = 0$ $\lambda_{1,2} = 2$ $\lambda_{1,3} = 1$	$\lambda_{2,1} = 0$ $\lambda_{2,2} = 0$ $\lambda_{2,3} = 0$ $\lambda_{2,4} = 1$	$\lambda_{3,1} = 1$ $\lambda_{3,2} = 0$ $\lambda_{3,3} = 0$ $\lambda_{3,4} = 2$
$F_1 = 1$	$F_2 = 3$	$F_3 =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

Figure 3.3. Notation example.

of each measurement: given a set of reports extracted from a single SAR image and a postulated 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.

The coarse-level uncertainty is inherent in the multiplicity of target primitives and distinct features extracted from each image. It is further compounded by the stochastic nature of the peak detection process, whereby some primitives will not generate reports in a given image (*i.e.*, missed detections) and some reports will correspond to no primitive (*i.e.*, false alarms), so that in general there will not be exhaustive correspondence between the target primitives and the reports in any image. To characterize this 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, assuming that any report is attributable to at most one primitive (an assumption which will be formalized shortly), we define a *label parameter* $\lambda_{k,j}$ describing the identity of each report $\mathbf{Z}_{k,j}$:

$$\lambda_{k,j} = \begin{cases} i, & \text{if report } \mathbf{Z}_{k,j} \text{ corresponds to target primitive } i, \\ 0, & \text{if report } \mathbf{Z}_{k,j} \text{ is spurious (corresponds to no primitive).} \end{cases} \tag{3.15}$$

The label parameter $\lambda_{k,j}$ corresponding to any image feature $\mathbf{Z}_{k,j}$ is, of course, unobservable, but offers a concise identification of that feature in terms of its generating primitive. For convenience in future expressions, we define F_k to be the number of false alarms in image k, *i.e.*, the number of $\lambda_{k,j}$ which equal 0 for a given k. Figure 3.3 presents an illustrative example of the notation and concepts encapsulated in $\lambda_{k,j}$. This figure depicts a simple scenario involving two target primitives ($N = 2$) and three images ($K = 3$).

It will be convenient to define a vector $\boldsymbol{\lambda}_k$ collecting the label parameters for all of the reports in image k:

$$\boldsymbol{\lambda}_k = \begin{bmatrix} \lambda_{k,1} \\ \vdots \\ \lambda_{k,M_k} \end{bmatrix}. \tag{3.16}$$

The vector $\boldsymbol{\lambda}$ introduced above can be formally defined as

$$\boldsymbol{\lambda} = \begin{bmatrix} \boldsymbol{\lambda}_1 \\ \vdots \\ \boldsymbol{\lambda}_K \end{bmatrix}. \tag{3.17}$$

The vector $\boldsymbol{\lambda}$ thus encompasses the identities of all features in all images. Although $\boldsymbol{\lambda}$ has a general structure imposed by constraints on each $\boldsymbol{\lambda}_k$ (*e.g.*, each $\boldsymbol{\lambda}_k$ can contain at most N nonzero elements and cannot contain the same nonzero index more than once), the set of allowable $\boldsymbol{\lambda}$ is vast: it is the product of the sets of allowable $\boldsymbol{\lambda}_k$ for $k = 1, \ldots, K$, each of which is a set whose size increases combinatorially with N and M_k. For example, if $M_k = N$ then the number of potential $\boldsymbol{\lambda}_k$, allowing for missed detections and false alarms, is

$$\sum_{n=0}^{N} \binom{N}{n} \binom{N}{N-n} (N-n)! = \sum_{n=0}^{N} \frac{(N!)^2}{(N-n)!(n!)^2}. \tag{3.18}$$

For a scenario with $M_k = N = 5$, (3.18) is greater than 1000; for $M_k = N = 6$, it is greater than 10,000. Thus in a scenario with as few as 10 images, each with $M_k = N = 6$, $\boldsymbol{\lambda}$ is drawn from a space of more than 10^{40} elements.

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

$$p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta}) = p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta})\, p(\boldsymbol{\lambda}|\boldsymbol{\theta}), \tag{3.19}$$

a product of the fine-level probability density function (pdf) and the coarse-level probability mass function (pmf).[7] As (3.19) indicates, the unknown correspondences $\boldsymbol{\lambda}$ must play a fundamental role in any procedure for estimation of $\boldsymbol{\theta}$ from $\mathbf{Z}$. Applications in which measured data have unknown identities that must be determined in order to process the data properly are generally known as *correspondence* or *data association* problems. Data association problems arise in many areas, including multitarget tracking [5,63], object recognition [19,110], and tomographic reconstruction [64,98].

[7] Throughout this thesis we will 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 $p(\cdot)$ in both cases.

We make a number of general assumptions about the relationship of $\boldsymbol{\lambda}$ and $\mathbf{Z}$ to $\boldsymbol{\theta}$ that facilitate the specification of a measurement model in the form of (3.19). The first three assumptions concern the coarse-level uncertainty expressed by $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$; the last two concern the remaining fine-scale uncertainty expressed by $p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta})$. All five of these assumptions are largely justifiable on simple physical grounds, and are standard in a variety of data association contexts. These assumptions are as follows:

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

Comments. This assumption is standard in nearly all data association contexts and is essentially a statement that the target model includes all significant scatterers.

Assumption 2. *The detectability of the ith 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.*

Comments. This is also a standard assumption, and in this context is a statement about the completeness of $\boldsymbol{\theta}_i$—namely, that a primitive's detectability is influenced only by its inherent description and is not influenced by other primitives or unmodeled effects. This assumption is violated by effects such as obstruction, in which a primitive's detectability might be adversely affected by the locations of other primitives. A relaxation of this assumption to accommodate obstruction will be examined in Chapter 6.

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

Comments. This assumption was inherent in the introduction of the label parameters in (3.15) and essentially states that the targets under consideration can be well-modeled as spatially distributed sets of discrete primitives—an assumption that is made implicitly by the parametric scattering representations of many of the efforts described in Section 2.4 (though not motivated by data association concerns as it is here). This is similar to assumptions made in many other data association contexts, including multitarget tracking [5, 63]. This assumption neglects scatterer-interaction effects such as multiple-primitive reflections. Chapter 7 includes a detailed analysis of a target from which multiple-primitive reflections are observed; the results of that chapter suggest that the algorithm is relatively robust to this type of primitive coupling. In Chapter 8 we examine possible approaches to dealing with scatterer-interaction effects explicitly.

Assumption 4. *Any report in any image corresponding to the ith primitive depends only on* $\boldsymbol{\theta}_i$*; furthermore, reports in a single image and between images are conditionally independent given* $\boldsymbol{\theta}$ *and* $\boldsymbol{\lambda}$*, whether they are detections or false alarms.*

Comments. Again, this is a standard assumption in most data association applications, and like Assumption 2 is a statement about the completeness of our parameterization. Phenomena that cannot be adequately described in the context of the chosen parameterization or by the assumed physical models could compromise this assumption. One such phenomenon—namely, responses observed from noncanonical primitives—is considered in Chapter 7.

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

Comments. This assumption, like Assumptions 2 and 4, can also be viewed as a statement about the completeness of $\boldsymbol{\theta}$. Its use is a bit of an idealization: for instance, we might expect bright responses to yield better $\mathbf{x}_{k,j}$ measurements, thus coupling $\mathbf{x}_{k,j}$ and $a_{k,j}$. As we will discuss in Chapter 8, however, the relaxation of this assumption introduces no conceptual changes to our algorithm. Chapters 6 and 7 include results from targets in which this assumption is violated.

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

$$p(\boldsymbol{\lambda}|\boldsymbol{\theta}) = \prod_{k=1}^{K} p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}). \tag{3.20}$$

Similarly, Assumption 4 implies that $p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta})$ can be factored as

$$p(\mathbf{Z}|\boldsymbol{\lambda},\boldsymbol{\theta}) = \prod_{k=1}^{K} p(\mathbf{Z}_k|\boldsymbol{\lambda}_k,\boldsymbol{\theta}) = \prod_{k=1}^{K} \left(\prod_{j=1}^{M_k} p(\mathbf{Z}_{k,j}|\lambda_{k,j},\boldsymbol{\theta}) \right). \tag{3.21}$$

Assumption 5 provides a further factorization:

$$p(\mathbf{Z}_{k,j}|\lambda_{k,j},\boldsymbol{\theta}) = p(\mathbf{x}_{k,j}|\lambda_{k,j},\boldsymbol{\theta})\, p(a_{k,j}|\lambda_{k,j},\boldsymbol{\theta})\, p(t_{k,j}|\lambda_{k,j},\boldsymbol{\theta}). \tag{3.22}$$

Although, as indicated in the comments above, there are situations in which these assumptions will fail, they are largely realistic and standard in similar problems. The last three chapters of this thesis are dedicated in large part to examining relaxations of these assumptions in order to accommodate a greater range of real-world effects.

To complete the measurement model of (3.19) we now need only specify the terms on the right-hand sides of (3.20) and (3.22). The framework provided by these equations is quite general, in the sense that they allow a wide variety of possible models for the dependence of $\mathbf{Z}$ on $\boldsymbol{\theta}$. In the following two sections we construct a specific measurement model that enables us to demonstrate our model generation technique in a concrete framework in subsequent chapters. A different choice of $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$ in (3.20) or of the pdfs on the right-hand side of (3.22) would lead to an algorithm only cosmetically different from the one presented here. We now specify the terms of (3.20) and (3.22) in turn.

3.3.1 Coarse-Level Uncertainty Model

The imagewise identity-uncertainty term $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$ required by (3.20) 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 and it cannot contain the same nonzero index twice. To facilitate specification of $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$, we introduce a vector of (unobservable) primitive-detection indicators

$$\boldsymbol{\delta}_k = \begin{bmatrix} \delta_{1,k} \\ \vdots \\ \delta_{N,k} \end{bmatrix}, \tag{3.23}$$

where each $\delta_{i,k}$ is a binary indicator signaling whether primitive i generated a report in image k, *i.e.*:

$$\delta_{i,k} = \begin{cases} 1, & \text{if } \lambda_{k,j} = i \text{ for some } j = 1, \ldots, M_k, \\ 0, & \text{if } \lambda_{k,j} \neq i \text{ for all } j = 1, \ldots, M_k. \end{cases} \tag{3.24}$$

(Thus the scenario of Figure 3.3 would give $\boldsymbol{\delta}_1 = \begin{bmatrix}1 & 1\end{bmatrix}^T$, $\boldsymbol{\delta}_2 = \begin{bmatrix}1 & 0\end{bmatrix}^T$, and $\boldsymbol{\delta}_3 = \begin{bmatrix}1 & 1\end{bmatrix}^T$.) We may then write

$$p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}) = p(\boldsymbol{\lambda}_k|F_k, \boldsymbol{\delta}_k, \boldsymbol{\theta})\, p(F_k, \boldsymbol{\delta}_k|\boldsymbol{\theta}), \tag{3.25}$$

which Assumptions 1 and 2 tell us is equivalent to

$$p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}) = p(\boldsymbol{\lambda}_k|F_k, \boldsymbol{\delta}_k, \boldsymbol{\theta})\, p(F_k) \prod_{i=1}^{N} p(\delta_{i,k}|\boldsymbol{\theta}_i). \tag{3.26}$$

Note that given F_k and $\boldsymbol{\delta}_k$ the only remaining uncertainty in $\boldsymbol{\lambda}_k$ is the ordering of its entries. If there is no systematic or preferred ordering, then each permutation is equally likely and we have

$$p(\boldsymbol{\lambda}_k|F_k, \boldsymbol{\delta}_k, \boldsymbol{\theta}) = \left(\binom{M_k}{F_k} (M_k - F_k)! \right)^{-1} = \frac{F_k!}{M_k!}. \tag{3.27}$$

For $p(F_k)$, we assume a standard Poisson false-alarm model that gives us the pmf

$$p(F_k) = \frac{e^{-\gamma_{FA} V} (\gamma_{FA} V)^{F_k}}{F_k!}, \tag{3.28}$$

where V denotes the sensor volume, γ_{FA} is a false alarm rate which may be estimated by running the feature extractor on characteristic imagery, and where $F_k \in \mathbb{Z}_+$ implicitly. To specify the third and final term of (3.26), we model a primitive's detectability as a function of its RCS at any viewing angle. (This function can be estimated by running the feature extractor on characteristic imagery containing scatterers of known amplitude.) We denote this probability-of-detection function as $P_D(.)$ and write

$$P_{D}{}'_{k,i} = P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k)), \tag{3.29}$$

so that

$$p(\delta_{i,k}|\boldsymbol{\theta}_i) = (P_{D}{}'_{k,i})^{\delta_{i,k}}(1 - P_{D}{}'_{k,i})^{1-\delta_{i,k}}. \tag{3.30}$$

The product of N of these terms can be expressed as

$$\prod_{i=1}^{N} p(\delta_{i,k}|\boldsymbol{\theta}_i) = \prod_{i=1}^{N} \left(1 - P_{D}{}'_{k,i}\right) \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P_{D}{}'_{k,\lambda_{k,j}}}{1 - P_{D}{}'_{k,\lambda_{k,j}}}. \tag{3.31}$$

We can now combine (3.27), (3.28), and (3.31) to obtain a complete model for the coarse-level uncertainty in image k:

$$p(\boldsymbol{\lambda}_k|\boldsymbol{\theta}) = \frac{e^{-\gamma_{FA}V}(\gamma_{FA}V)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} \left(1 - P_{D}{}'_{k,i}\right) \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P_{D}{}'_{k,\lambda_{k,j}}}{1 - P_{D}{}'_{k,\lambda_{k,j}}}. \tag{3.32}$$

The overall coarse-level pdf $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$ is simply the product of K of these terms as in (3.20).

■ 3.3.2 Fine-Level Uncertainty Model

Specification of the distributions of report parameters in (3.22) requires consideration of two cases: one in which the report under consideration is a detection (*i.e.*, $\lambda_{k,j} \neq 0$), and one in which it is a false alarm (*i.e.*, $\lambda_{k,j} = 0$). We model false alarms as being uniformly distributed throughout a SAR image, equally likely to be classified as an even- or odd-bounce scatterer, and having amplitudes that are tabulated in an empirical pdf $p_{FA}(.)$ (obtained, like γ_{FA} and $P_D(.)$, by processing characteristic imagery), so that

$$p(\mathbf{x}_{k,j}|\lambda_{k,j} = 0, \boldsymbol{\theta}) = \frac{1}{V}, \tag{3.33}$$

$$p(t_{k,j}|\lambda_{k,j} = 0, \boldsymbol{\theta}) = \frac{1}{2}, \tag{3.34}$$

and

$$p(a_{k,j}|\lambda_{k,j} = 0, \boldsymbol{\theta}) = p_{FA}(a_{k,j}). \tag{3.35}$$

For a report corresponding to a detected primitive, the physical and imaging models of Chapter 2 provide a basis for the form of each term in (3.22). In particular, $\mathbf{x}_{k,j}$ can be considered a measurement of $\boldsymbol{\pi}_k\big(\boldsymbol{\theta}_{\lambda_{k,j}}\big)$ in (3.3), and we model

$$p(\mathbf{x}_{k,j}|\lambda_{k,j} \neq 0, \boldsymbol{\theta}) = \frac{\exp\left[-\frac{1}{2}\left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k\big(\boldsymbol{\theta}_{\lambda_{k,j}}\big)\right)^T \mathbf{R}^{-1}\left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k\big(\boldsymbol{\theta}_{\lambda_{k,j}}\big)\right)\right]}{2\pi(\det \mathbf{R})^{1/2}} \tag{3.36}$$

(*i.e.*, as a Gaussian with mean $\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}})$ and covariance $\mathbf{R}$), where $\mathbf{R}$ can be estimated from characteristic imagery. Additionally, if each SAR image is narrowband and narrow-aperture (so that a scatterer's RCS throughout the phase history shows little variation from its value at the center frequency and center viewing angle), then $a_{k,j}$ will be a good measurement of $A(\boldsymbol{\theta}_{\lambda_{k,j}}, \psi_k, \phi_k)$ in (3.5), and we model

$$p(a_{k,j}|\lambda_{k,j} \neq 0, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma_a^2)^{1/2}} \exp\left[-\frac{1}{2\sigma_a^2}(a_{k,j} - A(\boldsymbol{\theta}_{\lambda_{k,j}}, \psi_k, \phi_k))^2\right], \tag{3.37}$$

where σ_a^2, like $\mathbf{R}$, can be estimated from characteristic imagery. Finally, to model $t_{k,j}$, we assume the availability of an $n_t \times 2$ confusion matrix $\boldsymbol{\rho}$ (obtained by processing characteristic imagery containing primitives of known type), such that

$$[\boldsymbol{\rho}]_{\theta^t_{\lambda_{k,j}}, t_{k,j}} = p(t_{k,j}|\lambda_{k,j} \neq 0, \boldsymbol{\theta}). \tag{3.38}$$

To simplify notation in subsequent expressions, we write

$$\rho'_{k,j} = [\boldsymbol{\rho}]_{\theta^t_{\lambda_{k,j}}, t_{k,j}} \tag{3.39}$$

for any detection. We can now combine (3.33)–(3.39) to yield the fine-level uncertainty model for image k:

$$\begin{aligned} p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta}) = {} & \left(\frac{1}{2V}\right)^{F_k} \cdot \prod_{j:\lambda_{k,j}=0} p_{FA}(a_{k,j}) \\ & \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{\exp\left(-\frac{1}{2\sigma_a^2}(a_{k,j} - A(\boldsymbol{\theta}_{\lambda_{k,j}}, \psi_k, \phi_k))^2\right)}{(2\pi\sigma_a^2)^{1/2}} \cdot \rho'_{k,j} \\ & \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) - \mathbf{x}_{k,j}\right)^T \mathbf{R}^{-1}\left(\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) - \mathbf{x}_{k,j}\right)\right)}{2\pi(\det \mathbf{R})^{1/2}}. \end{aligned} \tag{3.40}$$

The overall fine-level pdf $p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta})$ is simply the product of K of these terms as in (3.21).

■ 3.3.3 Complete Measurement Model and Formal Estimation Statement

We now have models for $p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta})$ and $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$ that can be combined to yield a complete measurement model $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$. This can be expressed as

$$p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta}) = \prod_{k=1}^{K} p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta}), \tag{3.41}$$

quantity	description
$\mathbf{R}$	2×2 covariance of $\mathbf{x}_{k,j}$ for detections
$\boldsymbol{\rho}$	$n_t \times 2$ confusion matrix of $t_{k,j}$ for detections
σ_a^2	variance of $a_{k,j}$ for detections
$P_D(\cdot)$	function mapping scatterer amplitude to probability of detection
$p_{FA}(\cdot)$	pdf of $a_{k,j}$ for false alarms
γ_{FA}	false alarm rate

Table 3.2. Measurement-model parameters and functions. These terms comprise a statistical characterization of the feature extraction process.

where

$$\begin{aligned} p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta}) = {} & \frac{e^{-\gamma_{FA} V} \left(\frac{\gamma_{FA}}{2}\right)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} (1 - P_D{}'_{k,i}) \cdot \prod_{j:\lambda_{k,j}=0} p_{FA}(a_{k,j}) \\ & \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P_D{}'_{k,\lambda_{k,j}}}{1 - P_D{}'_{k,\lambda_{k,j}}} \cdot \frac{\exp\left(-\frac{1}{2\sigma_a^2}(a_{k,j} - A(\boldsymbol{\theta}_{\lambda_{k,j}}, \psi_k, \phi_k))^2\right)}{(2\pi\sigma_a^2)^{1/2}} \cdot \rho'_{k,j} \quad (3.42) \\ & \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{\exp\left(-\frac{1}{2}\left(\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) - \mathbf{x}_{k,j}\right)^T \mathbf{R}^{-1} \left(\boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) - \mathbf{x}_{k,j}\right)\right)}{2\pi(\det \mathbf{R})^{1/2}}. \end{aligned}$$

Table 3.2 summarizes the parameters and functions of (3.42) that must be specified by estimation from characteristic imagery or by other means.

The measurement model of (3.41) and (3.42) is a complete description of the probabilistic relationship of extracted features $\mathbf{Z}$ to target parameters $\boldsymbol{\theta}$. Our goal is the maximum likelihood (ML) estimation of $\boldsymbol{\theta}$ from $\mathbf{Z}$, *i.e.*,

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\mathbf{Z}|\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} \sum_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta}), \quad (3.43)$$

where $\boldsymbol{\Lambda}$ is the set of all possible $\boldsymbol{\lambda}$. The chief difficulties in obtaining such an estimate are the vast size of $\boldsymbol{\Lambda}$ and the structure that the summation over $\boldsymbol{\Lambda}$ imparts to $p(\mathbf{Z}|\boldsymbol{\theta})$. In particular, $p(\mathbf{Z}|\boldsymbol{\theta})$ is a high-dimensional mixture density fraught with local maxima. Any attempt to obtain $\hat{\boldsymbol{\theta}}$ as in (3.43) must reckon with these fundamental difficulties. This is the focus of the next chapter.

Chapter 4

An Expectation-Maximization Approach to Model Generation

THE measurement model of (3.42) 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 vehicle 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, with $p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta})$ providing a conditional likelihood function that could be maximized by standard means to produce an ML estimate [97, 105]. The unobservability of $\boldsymbol{\lambda}$, however, greatly complicates the problem.

The chief difficulty facing almost all data association problems [5, 19, 63, 64, 98, 110], including the one described here, is the combinatorial proliferation of possible correspondences that make direct maximization of the marginal likelihood function $p(\mathbf{Z}|\boldsymbol{\theta})$ intractable. One way to manage the combinatorial explosion of possibilities is to dismiss as infeasible a majority of associations corresponding to extremely unlikely events; we will utilize a technique known as *gating*, to be described later, for this purpose [5]. 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 [28, 71, 75, 91]. 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.

■ 4.1 The Expectation-Maximization Method

The EM method, first proposed in 1977 by Dempster, Laird, and Rubin [28], is an iterative procedure for producing an ML estimate of parameters in incomplete-data problems, *i.e.*, when there is a many-to-one mapping from a postulated set of "complete" data to the set of observed data. The general incomplete-data problem is as follows. An ML estimate of $\boldsymbol{\Psi}$ is sought from a vector of *observed data* $\mathbf{Y}$:

$$\hat{\boldsymbol{\Psi}}_{\mathrm{ML}} = \arg\max_{\boldsymbol{\Psi}} \log p(\mathbf{Y}|\boldsymbol{\Psi}). \tag{4.1}$$

The maximization of (4.1) is often difficult to achieve, but might be simpler if $\mathbf{Y}$ could be augmented with a vector of *unobserved data* $\mathbf{X}$. The vectors $\mathbf{X}$ and $\mathbf{Y}$ together form the *complete data*, distributed according to $p(\mathbf{X}, \mathbf{Y}|\boldsymbol{\Psi})$. There is said to be a many-to-one mapping from the set of complete data to the observed data because in general $\mathbf{X}$ might take on any one of a number of values for a given observation $\mathbf{Y}$.

The EM method achieves the maximization of (4.1) by successive maximizations of an *expected-log-likelihood* function $Q(.)$ that is related to $p(\mathbf{Y}|\boldsymbol{\Psi})$ but is generally easier to maximize. Each iteration of the EM method consists of two steps: an expectation (E) step and a maximization (M) step. The E step entails calculation of

$$\begin{aligned} Q(\boldsymbol{\Psi}|\boldsymbol{\Psi}^{[n]}) &= E\left[\log p(\mathbf{X}, \mathbf{Y}|\boldsymbol{\Psi}) \,\middle|\, \mathbf{Y}, \boldsymbol{\Psi}^{[n]}\right] \\ &= \int_{\boldsymbol{\mathcal{X}}} \left[\log p(\mathbf{X}, \mathbf{Y}|\boldsymbol{\Psi})\right] p(\mathbf{X}|\mathbf{Y}, \boldsymbol{\Psi}^{[n]})\, d\mathbf{X}, \end{aligned} \tag{4.2}$$

where $\boldsymbol{\Psi}^{[n]}$ is the estimate of $\boldsymbol{\Psi}$ from the nth iteration and $\boldsymbol{\mathcal{X}}$ is the space from which $\mathbf{X}$ is drawn. The M step then maximizes this quantity to obtain a new estimate of $\boldsymbol{\Psi}$:

$$\boldsymbol{\Psi}^{[n+1]} = \arg\max_{\boldsymbol{\Psi}} Q(\boldsymbol{\Psi}|\boldsymbol{\Psi}^{[n]}). \tag{4.3}$$

(Figure 4.1 is a conceptual representation of the EM method.) The iteration between the E and M steps is essentially a bootstrap approach to generating an estimate of $\boldsymbol{\Psi}$ given little initial knowledge of the hidden data $\mathbf{X}$: as $\boldsymbol{\Psi}^{[n]}$ improves, the E-step calculation of $Q(\boldsymbol{\Psi}|\boldsymbol{\Psi}^{[n]})$ provides a better local approximation to the true log-likelihood function, and as $Q(\boldsymbol{\Psi}|\boldsymbol{\Psi}^{[n]})$ improves, the M-step maximization yields a better estimate $\boldsymbol{\Psi}^{[n]}$. In fact, one property that makes the EM method quite attractive is the guarantee that the likelihood of successive estimates will never decrease [28], *i.e.*,

$$p(\mathbf{Y}|\boldsymbol{\Psi}^{[n+1]}) \geq p(\mathbf{Y}|\boldsymbol{\Psi}^{[n]}). \tag{4.4}$$

Thus, as long as the likelihood function is bounded, the iteration will result in convergence to at least a local maximum in the likelihood function.[1]

Note that the EM method does not provide any indication of how to carry out the expectation of (4.2) or the maximization of (4.3). The precise forms of these equations will depend on the problem at hand, and there is no guarantee that either step will be easy. The calculation of $p(\mathbf{X}|\mathbf{Y}, \boldsymbol{\Psi}^{[n]})$ in the E step might be intractable, or $Q(\boldsymbol{\Psi}|\boldsymbol{\Psi}^{[n]})$ might be highly nonlinear and difficult to maximize in the M step. Selection of an initial estimate $\boldsymbol{\Psi}^{[0]}$ (required to begin the iteration) might be problematic. In short, the relative simplicity of (4.2) and (4.3) often belies the sophisticated machinery required to implement the EM method for a given problem.

The remaining sections of this chapter describe an EM-based algorithm for target model generation (*i.e.*, for implementing the "model generation" stage of Figure 3.1). In this context, the observed data is $\mathbf{Z}$ and the hidden data is $\boldsymbol{\lambda}$. These quantities

[1]Further convergence properties of the EM method are examined in [11, 71, 117].

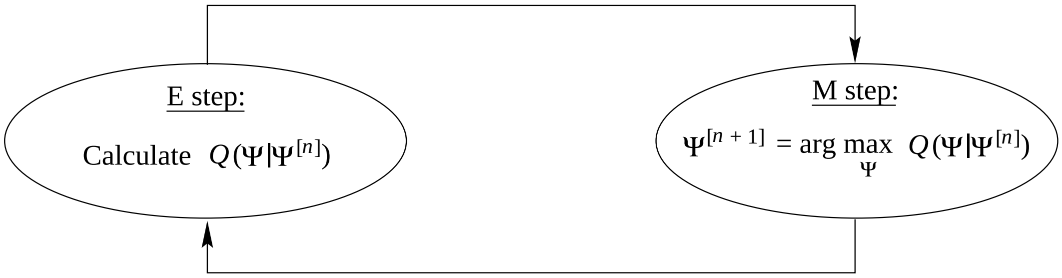


Figure 4.1. General form of the EM method.

are related to $\boldsymbol{\theta}$ according to the measurement model $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$, which determines the expected-log-likelihood function $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$. A block diagram of our basic algorithm is depicted in Figure 4.2. The E and M steps that form the core of this algorithm are described in Sections 4.2 and 4.3, respectively. Because the EM method is an iterative procedure, it requires an initialization $\boldsymbol{\theta}^{[0]}$; our procedure for producing such an initialization is described in Section 4.4. 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 E and M steps we develop in Sections 4.2 and 4.3 assume that N is specified. The initialization stage and a model-order adjustment stage, described in Section 4.5, together enable adaptive selection of model order as the algorithm progresses. Finally, the EM method requires a criterion for determining convergence and terminating the iteration. Our termination criterion is described in Section 4.6.

■ 4.2 The E Step

In our framework the form of the E step of the algorithm can be stated in specific terms as

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) &= E\left[\log p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta}) \;\middle|\; \mathbf{Z}, \boldsymbol{\theta}^{[n-1]}\right] \\ &= \sum_{\boldsymbol{\lambda} \in \boldsymbol{\Lambda}} \left[\log p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})\right] p(\boldsymbol{\lambda}|\mathbf{Z}, \boldsymbol{\theta}^{[n-1]}), \end{aligned} \tag{4.5}$$

where $\boldsymbol{\Lambda}$ is the set of all possible $\boldsymbol{\lambda}$. The derivation of Appendix B shows that with N specified, (4.5) can be written as

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_{i=1}^{N} Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) + C_K, \tag{4.6}$$

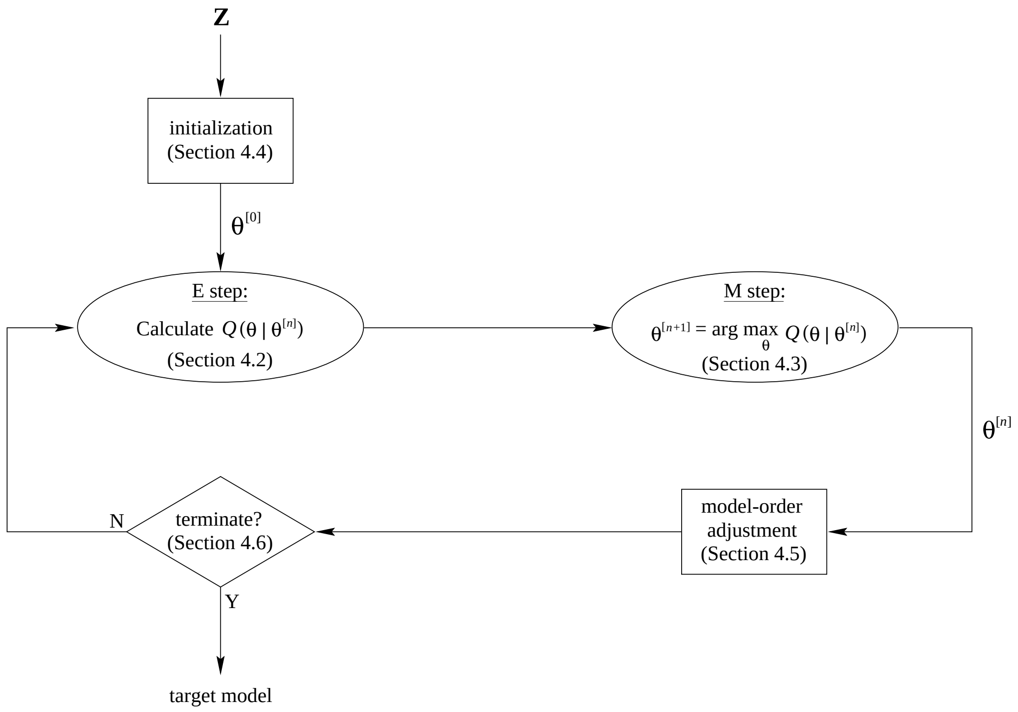


Figure 4.2. Implementation of an EM-based model-generation algorithm.

where each of the N terms $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ depends only on primitive i and can be further decomposed into a sum of K terms, one for each image:

$$Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{k=1}^{K} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}), \tag{4.7}$$

where

$$\begin{aligned} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \Bigg[& \log \frac{P_{D'_{k,i}}}{1 - P_{D'_{k,i}}} - \frac{1}{2\sigma_a^2} \left(a_{k,j} - A(\boldsymbol{\theta}_i, \psi_k, \phi_k)\right)^2 \\ & + \log \rho'_{k,j} - \frac{1}{2} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)^T \mathbf{R}^{-1} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right) \Bigg] \\ & + \log(1 - P_{D'_{k,i}}), \end{aligned} \tag{4.8}$$

or, equivalently,

$$
\begin{aligned}
Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \Pr(\delta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \log P_{D'_{k,i}} + \Pr(\delta_{i,k} = 0|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \log \left(1 - P_{D'_{k,i}}\right) \\
+ \sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \Bigg[-\frac{1}{2\sigma_a^2} \left(a_{k,j} - A(\boldsymbol{\theta}_i, \psi_k, \phi_k)\right)^2 \\
+ \log \rho'_{k,j} - \frac{1}{2} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)^T \mathbf{R}^{-1} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right) \Bigg].
\end{aligned}
\tag{4.9}
$$

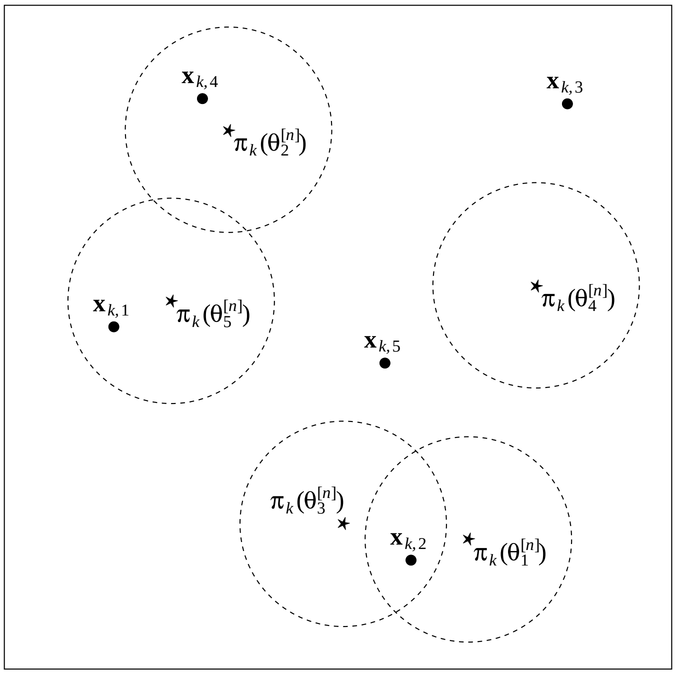
In other words, the expected-log-likelihood $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ separates into NK terms, each depending only on a single primitive and on the reports in a single image. This decoupling is a consequence of our independence assumptions of Section 3.3. (A similar decomposition will be possible in the M step.) As suggested by the form of (4.9), each $Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ can be viewed as a combination of $M_k + 1$ different penalties associated with primitive i in image k: a detection component penalizing any mismatch between the primitive's actual presence in image k and its predicted detectability in that image, and M_k components each penalizing the mismatch between the components of a single report in image k and the primitive's predicted location, amplitude, and type in that image. Each of the M_k report penalties is weighted by a *report-to-primitive correspondence probability* $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ that remains to be calculated.

Implementation of the E step requires a total of $N \sum_{k=1}^{K} M_k$ correspondence probabilities to be calculated—one for each (i, j, k) triplet. In theory this can be accomplished via Bayes' rule:

$$
\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) = \frac{\sum_{\boldsymbol{\lambda}_k : \lambda_{k,j} = i} p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}{\sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}, \tag{4.10}
$$

where $\boldsymbol{\Lambda}_k$ is the set of allowable $\boldsymbol{\lambda}_k$. Note that the assumptions of Section 3.3 greatly simplify the calculation of the report-to-primitive correspondence probabilities by requiring sums over only $\boldsymbol{\Lambda}_k$ instead of the much larger set $\boldsymbol{\Lambda}$. Even with this simplification, however, the computation of (4.10) is typically intractable even for problems of modest size due to the combinatorial dependence of the size of $\boldsymbol{\Lambda}_k$ on N and M_k. As indicated previously, we overcome this difficulty by gating [5, 63], a common and easily justifiable simplification. Specifically, complete enumeration of $\boldsymbol{\Lambda}_k$ 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 a set of feasible associations $\tilde{\boldsymbol{\Lambda}}_k^{[n]} \subset \boldsymbol{\Lambda}_k$ at each iteration, where

$$
\tilde{\boldsymbol{\Lambda}}_k^{[n]} = \left\{ \boldsymbol{\lambda}_k : \left\| \mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}^{[n]}) \right\| \leq r_{\text{gate}}, \quad j = 1, \ldots, M_k \right\}, \tag{4.11}
$$



Elements of $\tilde{\mathbf{\Lambda}}_k^{[n]}$
[5 1 0 2 0]
[5 1 0 0 0]
[5 3 0 2 0]
[5 3 0 0 0]
[5 0 0 2 0]
[5 0 0 0 0]
[0 1 0 2 0]
[0 1 0 0 0]
[0 3 0 2 0]
[0 3 0 0 0]
[0 0 0 2 0]
[0 0 0 0 0]

image k

Figure 4.3. Illustration of gating. Each primitive estimate at iteration n defines a circular gate region of radius r_{gate} centered at $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i^{[n]})$, the predicted location of primitive i in image k as in (3.3). Gating defines a set of feasible associations $\tilde{\mathbf{\Lambda}}_k^{[n]}$ that is limited to those $\boldsymbol{\lambda}_k \in \mathbf{\Lambda}_k$ for which each report $\mathbf{x}_{k,j}$ lies within the gate region of primitive $\lambda_{k,j}$, as in (4.11). In this illustration, $N = M_k = 5$, so (3.18) indicates that $\mathbf{\Lambda}_k$ contains 1546 elements; the set of feasible associations, however, consists of only twelve $\boldsymbol{\lambda}_k$.

and then making the assumption that

$$p(\boldsymbol{\lambda}_k, \mathbf{Z}_k | \boldsymbol{\theta}^{[n]}) = 0, \quad \boldsymbol{\lambda}_k \in \mathbf{\Lambda}_k \setminus \tilde{\mathbf{\Lambda}}_k^{[n]}. \tag{4.12}$$

(Typically r_{gate} is taken as a small multiple of $(\operatorname{trace} \mathbf{R})^{1/2}$.) Figure 4.3 is an illustration of how gating produces a set of feasible associations $\tilde{\mathbf{\Lambda}}_k^{[n]}$ that is much smaller than the original set $\mathbf{\Lambda}_k$. The effect of gating is thus a dramatic reduction in the number of $p(\boldsymbol{\lambda}_k, \mathbf{Z}_k | \boldsymbol{\theta}^{[n]})$ that must be calculated to obtain $\Pr(\lambda_{k,j} = i | \mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ as in (4.10). This makes the E step feasible. Gating has the added benefit of enabling non-overlapping gate regions to be considered independently in the calculation of the correspondence probabilities, further simplifying the E step.

A necessity for beneficial application of gating is that the current primitive location estimates $\theta_i^{\mathbf{x}[n]}$ must be reasonable approximations to the true primitive locations $\theta_i^{\mathbf{x}}$, or else $\tilde{\mathbf{\Lambda}}_k^{[n]}$ will likely exclude true report-to-primitive correspondences.[2] (This emphasizes

[2]The likelihood of this eventuality can be reduced by choosing a large value for r_{gate}. However, this

the necessity for a good initial estimate $\boldsymbol{\theta}^{[0]}$.) Assuming, however, that the location estimates at iteration n are sufficiently close to the $\theta_i^{\mathbf{x}}$ for the correct reports to be included in each gate region, gating has the desired effect of dismissing only extremely improbable report-to-primitive correspondences from consideration.

Before proceeding, we note that although the precise form of $Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ in (4.8) depends on the precise form of $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$, the basic decomposition of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ as in (4.6) and (4.7) would be possible for any measurement model, given the assumptions of Section 3.3. In particular, a different choice of measurement model would produce an expected-log-likelihood that is cosmetically different from the one here—but that would still decouple into NK separate terms and would still entail calculation of the $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$.

4.3 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 $\boldsymbol{\theta}$, *i.e.*,

$$\boldsymbol{\theta}^{[n+1]} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}). \tag{4.13}$$

The separation of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ into individual terms for each primitive in (4.6), made possible by the independence assumptions of Section 3.3, implies that the maximization of (4.13) 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 $\boldsymbol{\theta}_i$:

$$\boldsymbol{\theta}_i^{[n+1]} = \arg\max_{\boldsymbol{\theta}_i} Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}). \tag{4.14}$$

Since $\boldsymbol{\theta}_i$ includes both continuous parameters ($\theta_i^{\mathbf{x}}$, $\theta_i^{\mathbf{p}}$, 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 candidate maximizations over the continuous components of $\boldsymbol{\theta}_i$, one for each possible value of θ_i^t. Examination of (4.8) reveals that each candidate 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 through $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$, and the pose and base amplitude are coupled through $P_{D'_{k,i}}$ and $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$.

As it turns out, we can circumvent these difficulties by utilizing a slight modification of the basic EM method that is known as *expectation-conditional-maximization* (ECM) [71, 72]. ECM's fundamental modification to the standard EM method is its decomposition of the M step's joint maximization into multiple simpler maximizations,

is done at the cost of including more unlikely associations in $\tilde{\boldsymbol{\Lambda}}_k^{[n]}$ in cases where the estimated primitive locations are accurate.

each over a partitioned subvector of the complete parameter vector. For instance, for each of the n_t candidate maximizations, we can partition $\boldsymbol{\theta}_i$ as

$$\boldsymbol{\theta}_i = \mathbf{P} \begin{bmatrix} \theta_i^t \\ \boldsymbol{\theta}_i(1) \\ \boldsymbol{\theta}_i(2) \\ \boldsymbol{\theta}_i(3) \end{bmatrix}, \tag{4.15}$$

where θ_i^t is set to a fixed value, where

$$\boldsymbol{\theta}_i(1) = \theta_i^{\mathbf{P}}, \tag{4.16}$$

$$\boldsymbol{\theta}_i(2) = \theta_i^{a}, \tag{4.17}$$

$$\boldsymbol{\theta}_i(3) = \begin{bmatrix} \theta_i^{\mathbf{x}} \\ \theta_i^{r} \end{bmatrix}, \tag{4.18}$$

and where $\mathbf{P}$ is a permutation matrix that serves to reproduce the defined ordering of the elements of $\boldsymbol{\theta}_i$ as in (3.2). ECM then allows us to form $\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}$, the maximizing $\boldsymbol{\theta}_i$ conditioned on a type of θ_i^t, as

$$\boldsymbol{\theta}_i^{[n+1|\theta_i^t]} = \mathbf{P} \begin{bmatrix} \theta_i^t \\ \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(1) \\ \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(2) \\ \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(3) \end{bmatrix}, \tag{4.19}$$

where each $\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(\boldsymbol{.})$ is a conditional maximization over a single partition only:

$$\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(1) = \arg\max_{\boldsymbol{\theta}_i(1)} Q_i\left(\boldsymbol{\theta}_i(1) \,\middle|\, \theta_i^t, \boldsymbol{\theta}_i^{[n]}(1), \boldsymbol{\theta}_i^{[n]}(2), \boldsymbol{\theta}_i^{[n]}(3)\right), \tag{4.20}$$

$$\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(2) = \arg\max_{\boldsymbol{\theta}_i(2)} Q_i\left(\boldsymbol{\theta}_i(2) \,\middle|\, \theta_i^t, \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(1), \boldsymbol{\theta}_i^{[n]}(2), \boldsymbol{\theta}_i^{[n]}(3)\right), \tag{4.21}$$

$$\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(3) = \arg\max_{\boldsymbol{\theta}_i(3)} Q_i\left(\boldsymbol{\theta}_i(3) \,\middle|\, \theta_i^t, \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(1), \boldsymbol{\theta}_i^{[n+1|\theta_i^t]}(2), \boldsymbol{\theta}_i^{[n]}(3)\right). \tag{4.22}$$

(The composite parameter vector $\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}$ is thus obtained by a process that can be viewed as a single-iteration coordinate ascent [8].) The M-step estimate $\boldsymbol{\theta}_i^{[n+1]}$ is then taken to be the vector from the set of the n_t type-conditional maximizers that achieves the greatest expected-log-likelihood:

$$\boldsymbol{\theta}_i^{[n+1]} = \arg \max_{\substack{\boldsymbol{\theta}_i^{[n+1|\theta_i^t]}: \\ \theta_i^t = 1,\ldots,n_t}} Q_i(\boldsymbol{\theta}_i^{[n+1|\theta_i^t]} | \boldsymbol{\theta}_i^{[n]}). \tag{4.23}$$

Although the $\boldsymbol{\theta}_i^{[n+1]}$ obtained in such a manner is *not* guaranteed to be a global maximizer of $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ as in the standard EM method, ECM shares EM's monotone convergence properties (*i.e.*, the observed-data log-likelihood is still guaranteed not to decrease at any iteration) [72]. Furthermore, although ECM convergence may require more iterations than EM convergence for the same likelihood function, each iteration might be significantly faster and thus the overall processing time required to produce a convergent estimate may be shorter for ECM than for EM [71].

The ECM approach greatly simplifies the M step by decoupling components of the objective function into simpler pieces. Examination of (4.7) and (4.8) shows that maximization over either $\boldsymbol{\theta}_i(1)$ or $\boldsymbol{\theta}_i(2)$ requires consideration of the function

$$\begin{aligned}\sum_{k=1}^{K}\Bigg(\sum_{j=1}^{M_k}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})&\left[\log\frac{P_D{}'_{k,i}}{1-P_D{}'_{k,i}}-\frac{1}{2\sigma_a^2}\left(a_{k,j}-A(\boldsymbol{\theta}_i,\psi_k,\phi_k)\right)^2\right]\\ &+\log(1-P_D{}'_{k,i})\Bigg),\end{aligned}\tag{4.24}$$

the only terms in $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ that depend on $\theta_i^{\mathbf{P}}$ and θ_i^a. This is a highly nonlinear function. The maximization of (4.24) over the first partition (*i.e.*, pose) is achieved utilizing a coarse-to-fine search. In particular, the coarsest-level search utilizes a grid of azimuth/elevation points spread across the viewsphere[3] and, if applicable, the third Euler angle (axial rotation). Each successive refinement utilizes a local grid centered at the maximizing point from the previous grid. The maximization of (4.24) over the second partition (*i.e.*, base amplitude) is also accomplished using a coarse-to-fine search, though in a single dimension.

Examination of (4.7) and (4.8) shows that maximization over $\boldsymbol{\theta}_i(3)$ (*i.e.*, location and, if applicable, radius of curvature) requires consideration of

$$\sum_{k=1}^{K}\sum_{j=1}^{M_k}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\left[-\frac{1}{2}\left(\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)^T\mathbf{R}^{-1}\left(\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)\right],\tag{4.25}$$

the only terms in $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ that depend on $\theta_i^{\mathbf{x}}$ and θ_i^r. Maximization of (4.25) to achieve estimates of $\theta_i^{\mathbf{x}}$ and θ_i^r is thus a weighted least-squares problem. Defining

$$\tilde{\mathbf{R}}_{i,j,k}^{-1}=\mathbf{R}^{-1}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\tag{4.26}$$

and

$$\tilde{\mathbf{H}}_{i,k}=\left[\begin{array}{c:c}\mathbf{H}_k & \begin{array}{c}-\cos\psi'_{i,k}\\ 0\end{array}\end{array}\right]\tag{4.27}$$

[3]The construction of a uniformly spaced grid of points on the sphere is generally impossible; in fact, even choosing a set of points with roughly equal spacing is nontrivial [22]. We utilize a technique described in [113] to generate a "Hammersley point set" on the sphere. This is a set of azimuth/elevation points that achieves nearly uniform spacing on the sphere.

(where $\mathbf{H}_k$ is the projection matrix defined in (3.4)) for brevity, we can use the definition of $\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$ in (3.3) to obtain a closed-form solution for the maximizer of (4.25):

$$\theta_i^{\mathbf{x}\,[n+1]} = \left(\sum_{k=1}^{K}\sum_{j=1}^{M_k} \mathbf{H}_k^T \tilde{\mathbf{R}}_{i,j,k}^{-1} \mathbf{H}_k\right)^{-1} \left(\sum_{k=1}^{K}\sum_{j=1}^{M_k} \mathbf{H}_k^T \tilde{\mathbf{R}}_{i,j,k}^{-1} \mathbf{x}_{k,j}\right), \quad \theta_i^t \in \{1,3\}, \tag{4.28}$$

$$\begin{bmatrix} \theta_i^{\mathbf{x}\,[n+1]} \\ \theta_i^{r\,[n+1]} \end{bmatrix} = \left(\sum_{k=1}^{K}\sum_{j=1}^{M_k} \tilde{\mathbf{H}}_{i,k}^T \tilde{\mathbf{R}}_{i,j,k}^{-1} \tilde{\mathbf{H}}_{i,k}\right)^{-1} \left(\sum_{k=1}^{K}\sum_{j=1}^{M_k} \tilde{\mathbf{H}}_{i,k}^T \tilde{\mathbf{R}}_{i,j,k}^{-1} \mathbf{x}_{k,j}\right), \quad \theta_i^t \in \{2,4\}. \tag{4.29}$$

(If the maximizing radius of (4.29) is negative, we set $\theta_i^{r\,[n+1]}$ to 0 and use the $\theta_i^{\mathbf{x}\,[n+1]}$ provided by (4.28).)

As in the E step, we note that usage of a different measurement model would produce an M step differing only in detail from the one here, as long as the independence assumptions of Section 3.3 still held. In particular, the independence assumptions of that section ensure that for any specific choice of $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$, the M step could still proceed as a series of N independent maximizations (each of which could be implemented as several partitioned ECM maximizations as described above).

4.4 Initialization Procedure

Because the EM method is iterative, it requires an initial estimate $\boldsymbol{\theta}^{[0]}$ to begin the iteration, *i.e.*, to enable the calculation of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ in the first E step. In some scenarios, such an initialization might be provided by prior information (*e.g.*, an existing reflector-primitive target model, or a rough model of a target constructed by a human image analyst), but in general this initialization must be obtained from the same set of data that is used in the subsequent EM iteration. In this section we describe our procedure for producing an initialization $\boldsymbol{\theta}^{[0]}$ from $\mathbf{Z}$. As indicated previously, this initialization procedure works in conjunction with a model-order adjustment stage (described in Section 4.5) to select the model order adaptively as the EM iteration progresses. This model-order adjustment occurs after every M step, before the termination 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 overestimation of N: any overfit can be corrected in subsequent iterations by the model-order adjustment stage, but any underfit is permanent.

In general, the final parameter estimate provided by the EM method is dependent on the initialization [71]. This is especially true in situations like ours, in which $\boldsymbol{\lambda}$ imparts a multimodal structure to the likelihood function $p(\mathbf{Z}|\boldsymbol{\theta})$. An effective initialization procedure must select a $\boldsymbol{\theta}^{[0]}$ in the vicinity of the maximizing mode. In other words, the initialization procedure must implicitly or explicitly deal with the problem of assigning correspondences between reports, since the true correspondences determine

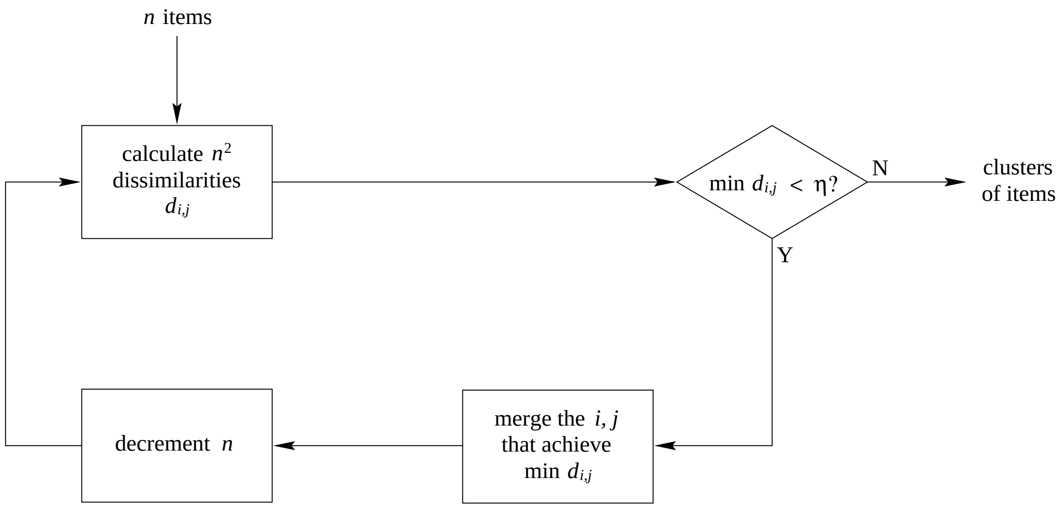


Figure 4.4. General form of agglomerative clustering.

the maximizing mode of $p(\mathbf{Z}|\boldsymbol{\theta})$. Our initialization procedure explicitly estimates these correspondences in order to construct $\boldsymbol{\theta}^{[0]}$. It is based on an agglomerative clustering algorithm that operates on the set of all $\mathbf{x}_{k,j}$ to group reports that are likely to have been generated by the same target primitive. 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 below, we can bias toward overfitting by setting an agglomeration threshold appropriately.

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 [1]. Figure 4.4 is a block diagram of the general agglomerative clustering process. 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 can 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. This notion can be made precise by examining the total squared error of the linear least-squares error (LLSE) estimate of the 3-D location of a hypothesized primitive responsible for a candidate group of reports. Specifically, suppose a candidate group $\boldsymbol{\mathcal{Z}}$ contains m reports:

$$\boldsymbol{\mathcal{Z}} = \{\mathbf{Z}_{k_1,j_1}, \ldots, \mathbf{Z}_{k_m,j_m}\}. \tag{4.30}$$

Suppose also that $\{k_1, \ldots, k_m\}$ has no repeated element, so the reports are from m different images. The LLSE location estimate for a fixed-specular-reflection-point primitive producing these m reports according to (3.3) and (3.36) is[4]

$$\hat{\mathbf{X}}_{\mathrm{L}}(\boldsymbol{\mathcal{Z}}) = \left(\sum_{i=1}^{m} \mathbf{H}_{k_i}^T \mathbf{R}^{-1} \mathbf{H}_{k_i}\right)^{-1} \left(\sum_{i=1}^{m} \mathbf{H}_{k_i}^T \mathbf{R}^{-1} \mathbf{x}_{k_i,j_i}\right). \tag{4.31}$$

We can write the normalized total squared error of this LLSE estimate as

$$\begin{aligned} \varepsilon_{\mathrm{L}}^2(\boldsymbol{\mathcal{Z}}) &= \sum_{i=1}^{m} \left\| \mathbf{H}_{k_i} \hat{\mathbf{X}}_{\mathrm{L}}(\boldsymbol{\mathcal{Z}}) - \mathbf{x}_{k_i,j_i} \right\|_{\mathbf{R}^{-1}}^2 \\ &= \sum_{i=1}^{m} (\mathbf{H}_{k_i} \hat{\mathbf{X}}_{\mathrm{L}}(\boldsymbol{\mathcal{Z}}) - \mathbf{x}_{k_i,j_i})^T \mathbf{R}^{-1} (\mathbf{H}_{k_i} \hat{\mathbf{X}}_{\mathrm{L}}(\boldsymbol{\mathcal{Z}}) - \mathbf{x}_{k_i,j_i}). \end{aligned} \tag{4.32}$$

It can be shown that if the m reports of $\boldsymbol{\mathcal{Z}}$ were in fact all produced by the same primitive according to the Gaussian model of (3.36), then $\varepsilon_{\mathrm{L}}^2(\boldsymbol{\mathcal{Z}})$ is a chi-squared random variable with $2m-3$ degrees of freedom.[5] On the other hand, if the m reports correspond to multiple primitives, then $\varepsilon_{\mathrm{L}}^2(\boldsymbol{\mathcal{Z}})$ will in general be much larger and will have a noncentral chi-squared distribution.[6] We thus use the chi-squared cumulative distribution function (cdf) as our measure of dissimilarity between clusters of reports. In particular, let $F_n(\cdot)$ be the cdf of a chi-squared random variable with n degrees of freedom, defined as

$$F_n(u) = \Pr(\chi_n^2 \leq u) = \int_0^u \frac{t^{(n/2)-1} e^{-t/2}}{2^{n/2}\Gamma(n/2)}\, dt. \tag{4.33}$$

Then for any two clusters of reports $\boldsymbol{\mathcal{Z}}_a$ and $\boldsymbol{\mathcal{Z}}_b$ that together comprise a set of reports

$$\boldsymbol{\mathcal{Z}}_a \cup \boldsymbol{\mathcal{Z}}_b = \{\mathbf{Z}_{k_1,j_1}, \ldots, \mathbf{Z}_{k_m,j_m}\}, \tag{4.34}$$

we define the dissimilarity between $\boldsymbol{\mathcal{Z}}_a$ and $\boldsymbol{\mathcal{Z}}_b$ to be

$$d(\boldsymbol{\mathcal{Z}}_a, \boldsymbol{\mathcal{Z}}_b) = \begin{cases} F_{2m-3}(\varepsilon_{\mathrm{L}}^2(\boldsymbol{\mathcal{Z}}_a \cup \boldsymbol{\mathcal{Z}}_b)), & k_i \neq k_j \text{ for all } i \neq j, \\ 1, & k_i = k_j \text{ for some } 1 \leq i, j \leq m. \end{cases} \tag{4.35}$$

[4] A similar expression can be written for radially symmetric primitives with wandering specular reflection points (*e.g.*, cylinders and tophats).

[5] This result follows from application of Theorem 9.9 of [95] to (4.32).

[6] If the reports include false alarms, which we model in (3.33) as having uniform rather than Gaussian distributions, then 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.

This provides an interpretation of $1-\eta$ as a confidence level (between 0 and 1) for accepting or rejecting a candidate group $\boldsymbol{\mathcal{Z}}$ based on $\varepsilon_{\mathrm{L}}^2(\boldsymbol{\mathcal{Z}})$. (Note that according to (4.35), clusters containing two reports from the same image are maximally dissimilar and will never be merged.) We would like to set η large enough so that most reports attributable to any single primitive will be grouped in a single cluster, but not so large as to group many reports from disparate primitives. In theory we could characterize such a chi-squared agglomeration process in terms of its probability of underfitting or overfitting a given target as a function of η: as η approaches 0, any two distinct clusters of reports are less likely to be agglomerated (even if they correspond to the same primitive) and the probability of underfitting goes to 0; as η approaches 1, distinct clusters of reports are more likely to be agglomerated (even if they correspond to distinct primitives) and the probability of underfitting goes to 1. Rather than attempt to characterize this variation completely, we set η to a value empirically determined to cluster most reports attributable to a single primitive while still giving 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 adjustment stage.

The chi-squared-statistic agglomerative clustering algorithm described here forms the heart of our initialization procedure. A block diagram of the complete initialization procedure is presented in Figure 4.5. A simple threshold test is first used to screen the complete set of reports $\mathbf{Z}$, rejecting those with low amplitudes (which are likely to be false alarms) and passing only those brighter than some threshold $a_{\min}$ to the chi-squared clusterer. The chi-squared agglomeration then proceeds as described above, producing a set of report clusters. Clusters containing only a small number of reports (say, two or three) are likely to be spurious and not necessarily indicative of a primitive; clusters containing fewer than some minimum number $n_{\min}$ of reports are thus removed at the conclusion of the agglomeration. Each remaining cluster is then used to initialize a single primitive vector $\boldsymbol{\theta}_i^{[0]}$ by means similar to those of Section 4.3. In particular, suppose that the ith cluster comprises reports $\boldsymbol{\mathcal{Z}}$ as defined in (4.30). We first initialize base amplitude to the maximum amplitude among $\boldsymbol{\mathcal{Z}}$:

$$\theta_i^{a\,[0]} = \max_{n=1,\ldots,m} a_{k_n,j_n} - c_{\mathrm{pol}}. \tag{4.36}$$

We then use the implicit estimate of $\boldsymbol{\lambda}$ provided by the grouping of $\boldsymbol{\mathcal{Z}}$ to define an initial set of estimated correspondence probabilities:

$$\Pr(\lambda_{k,j} = i | \mathbf{Z}_k, \boldsymbol{\theta}^{[-1]}) = \begin{cases} 1, & \mathbf{Z}_{k,j} \in \boldsymbol{\mathcal{Z}}, \\ 0, & \mathbf{Z}_{k,j} \notin \boldsymbol{\mathcal{Z}}. \end{cases} \tag{4.37}$$

These correspondence probabilities are then used to obtain the type, pose, location, and radius components of $\boldsymbol{\theta}_i^{[0]}$ by means identical to those described in Section 4.3, *i.e.*, by n_t candidate maximizations of (4.24) and (4.25) using the correspondence probabilities of (4.37).[7]

[7]Note that a location estimate is available as a byproduct of the chi-squared-statistic clustering; this

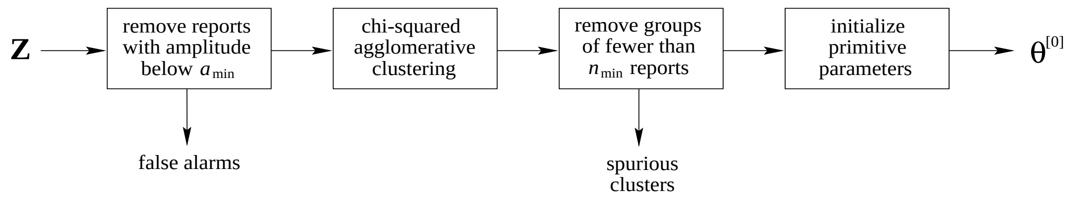


Figure 4.5. Block diagram of initialization procedure.

Before proceeding, we make two observations concerning the agglomerative clustering process. First, the agglomeration described here is based solely on report location, though it can easily be extended to include type and amplitude information. Second, note that the procedure we have described is not perfectly designed for grouping reports produced by cylinders and tophats since, as is evident from (3.3), the apparent specular 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 a tophat or cylinder 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 (see the results in Chapter 5).

■ 4.5 Model-Order Adjustment

Our model-order adjustment stage counteracts the overfitting induced in the initialization by gradually removing extraneous primitives. 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 $\tilde{p}_i^d$ for each primitive i for which $\boldsymbol{\theta}_i^{[n]}$ has converged:

$$\begin{aligned} \tilde{p}_i^d &= \frac{1}{K}\sum_{k=1}^{K} \Pr(\delta_{i,k} = 1 | \mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \\ &= \frac{1}{K}\sum_{k=1}^{K}\sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i | \mathbf{Z}_k, \boldsymbol{\theta}^{[n]}). \end{aligned} \tag{4.38}$$

(The $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ terms are available from the E step of the iteration.) Intuitively, if $\tilde{p}_i^d$ is near zero, then $\mathbf{Z}$ provides little evidence to support the hypothesis

estimate is the same as that provided by (4.28) using the correspondence probabilities of (4.37). For the cylinder and tophat candidate maximizations, (4.29) must be used to obtain a new location estimate along with a radius of curvature estimate.

of a target primitive whose parameters are as 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 a converged primitive's calculated $\tilde{p}_i^d$ 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 broader responses). This model-order adjustment stage ensures that all primitives contained in the final estimate of $\boldsymbol{\theta}$ 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 the hypothesized model order is still artificially large. The benefit of this extra computation is the increased adaptivity and robustness of the final estimate $\boldsymbol{\theta}$.

■ 4.6 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. In particular, at the conclusion of iteration n (see Figure 4.2), the algorithm is terminated if the model order was not adjusted between iterations $n-1$ and n, and if each $\boldsymbol{\theta}_i^{[n]}$ satisfies the following conditions:

$$\theta_i^{t\,[n]} = \theta_i^{t\,[n-1]}, \tag{4.39}$$

$$\left\| \theta_i^{\mathbf{x}\,[n]} - \theta_i^{\mathbf{x}\,[n-1]} \right\| \leq \tau_{\mathbf{x}}, \tag{4.40}$$

$$\left| \theta_i^{r\,[n]} - \theta_i^{r\,[n-1]} \right| \leq \tau_r, \tag{4.41}$$

$$\left| \theta_i^{a\,[n]} - \theta_i^{a\,[n-1]} \right| \leq \tau_a, \tag{4.42}$$

$$\sphericalangle(\theta_i^{\mathbf{p}\,[n]}, \theta_i^{\mathbf{p}\,[n-1]}) \leq \tau_{\mathbf{p}_1}, \tag{4.43}$$

$$\mathrm{rot}(\theta_i^{\mathbf{p}\,[n]}, \theta_i^{\mathbf{p}\,[n-1]}) \leq \tau_{\mathbf{p}_2}, \tag{4.44}$$

where each $\tau_{\bullet}$ is a small positive tolerance, $\sphericalangle(\bullet,\bullet)$ is the angular separation on the viewsphere between two Euler-angle poses, and $\mathrm{rot}(\bullet,\bullet)$ is the absolute difference between the rotation-angle components of two Euler-angle poses. If any $\boldsymbol{\theta}_i^{[n]}$ fails to satisfy any of (4.39)–(4.44), or if the model order was adjusted in the current iteration, then the termination criterion fails and the algorithm progresses to iteration $n+1$.

■ 4.7 Other Implementation Details

The previous sections give a complete overview of our basic algorithm for estimating target models from SAR data. This section highlights a few issues that are, conceptually,

stage	dominant contributor	approximate complexity
initialization	chi-squared clustering	$\mathcal{O}(K^2\bar{M}^2)$
E step	correspondence probability calculation	$\sim \mathcal{O}(NK\bar{M})$ for small r_{gate}
M step	pose maximization	$\mathcal{O}(NKn_t)$
model-order adjustment	—	$\mathcal{O}(NK\bar{M})$
termination check	—	$\mathcal{O}(N)$

Table 4.1. Complexity of each stage of the model generation algorithm. The second column lists the component that dominates each stage's computational burden. (The model-order adjustment stage and the termination check are computationally very simple and are essentially single-component stages.) The quantity $\bar{M}$ that appears in several terms of the third column is defined as the average number of reports in each image. Note that the E-step burden depends on r_{gate}.

of secondary importance, but are nevertheless relevant to the implementation of the algorithm as a computational procedure.

■ 4.7.1 Computational Burden

The implementation of the algorithm described in this chapter is made feasible by the independence assumptions of Section 3.3. In particular, these assumptions decouple $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ to the point that the computational complexity of each EM iteration is roughly linear in K and N. The complexity of each individual stage of the algorithm (see Figure 4.2) is summarized in Table 4.1. We now examine the relative burdens of these stages in closer detail.

The computational burden of the initialization stage is dominated by the agglomerative chi-squared clustering. Conceptually, this agglomeration requires calculation of LLSE location estimates and associated squared errors for all pairs of report clusters at all iterations of the agglomeration. This suggests that if N_R reports survive the pre-agglomeration amplitude thresholding (see Figure 4.5), then the computational burden of each agglomeration iteration is roughly quadratic in N_R. Furthermore, if the original set of N_R reports is eventually grouped into a small number of clusters, the overall agglomeration will require nearly N_R iterations. We might thus expect the overall computational complexity of the agglomeration to be cubic in N_R. In practice, a significant simplification is possible: many of the calculations required at any iteration will already have been performed in previous iterations, and each iteration will require only on the order of N_R new calculations of LLSE location estimates and associated squared errors. The overall complexity of the initialization stage is thus quadratic in N_R. Because N_R is roughly linearly proportional to K and $\bar{M}$ (the average number of reports in each image), the computational burden of the initialization stage is quadratic in K and $\bar{M}$.

As described in Section 4.2, the independence assumptions of Section 3.3 greatly reduce the number of calculations required to implement the E step by enabling the report-to-primitive correspondence probabilities to be calculated independently for each image. The E step complexity is thus linear in K. Gating provides a further simplification within each image by significantly reducing the number of feasible associations and enabling nonoverlapping gate regions (*i.e.*, gate regions that have no reports in common) to be considered independently. The complexity of the computation of all correspondence probabilities for a single image k depends strongly on r_{gate}, the gate radius. In the limiting case as $r_{\text{gate}} \to \infty$, we will have $\tilde{\boldsymbol{\Lambda}}_k^{[n]} \approx \boldsymbol{\Lambda}_k$ (see Section 4.2) and the computational burden will thus explode combinatorially as N and M_k increase. In the case where gate regions are nonoverlapping this burden is merely linear in N and M_k. In practice, for reasonably sized r_{gate}, most gate regions do not overlap and thus the E-step burden is generally nearly linear in N and $\bar{M}$, in addition to K.

The independence assumptions of Section 3.3 also simplify the M step considerably, allowing us to obtain $\boldsymbol{\theta}^{[n]}$ as N independent maximizations, each requiring n_t candidate maximizations over a single $\boldsymbol{\theta}_i$. The M step is thus linear in N and n_t.[8] The most computationally demanding component of each candidate $\boldsymbol{\theta}_i$ maximization is the partitioned pose maximization, which is implemented as a coarse-to-fine search in either two Euler angles (for cylinders and tophats) or three Euler angles (for dihedrals and trihedrals). Evaluation of the objective function (4.24) at each point in this search requires calculation of $S_{\theta_i^t}(\cdot)$ and other terms at K different viewing angles. The complicated nature of the $S_{\theta_i^t}(\cdot)$ (see Section 3.1 and Appendix A) makes this a nontrivial task. The pose maximization takes much longer than the other components of the M step (*i.e.*, the closed form maximization over location and radius and the one-dimensional coarse-to-fine search over base amplitude) and adds a linear dependence on K to the M-step complexity. In practice, the M step tends to dominate the computational burden of the overall algorithm.[9]

Compared to the initialization, E step, and M step, the termination check and model-order adjustment stage are inconsequential to the overall computational burden of the algorithm. The model-order adjustment stage has a complexity that is linear in N and K, whereas the termination check has a complexity that is linear in N and constant with respect to K, but neither stage adds significantly to the overall complexity of the model generation algorithm.

■ 4.7.2 Convergence Behavior

Although the likelihood attained by successive M step estimates $\boldsymbol{\theta}^{[n]}$ is guaranteed not to decrease at any iteration, there is in general no guarantee on the speed of convergence [71]. In practice, due to the separable nature of our $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$, we tend to observe convergence of some $\boldsymbol{\theta}_i$ before others. Thus while most $\boldsymbol{\theta}_i$ might converge to station-

[8] In Chapter 7 we present a modification to the basic algorithm that effectively reduces the M step's computational dependence on n_t to a constant.

[9] To mitigate the effects of the pose maximization on the overall computational burden of the algorithm, we coarsen the initial search grid once a primitive reaches convergence, *i.e.*, when the estimate of θ_i has reached a mode and is unlikely to change. Thus the M step will proceed more quickly as more primitive estimates converge.

ary values after only a few iterations, a single nonconvergent $\boldsymbol{\theta}_i$ can greatly prolong the duration of the overall estimation, since the EM cycle requires calculation of all $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ and maximization over all $\boldsymbol{\theta}_i$ at each iteration regardless of how many of the $\boldsymbol{\theta}_i$ have converged.

Empirical observation shows that as $\boldsymbol{\theta}_i^{[n]}$ converges, its associated report-to-primitive correspondence probabilities $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ tend to 0 and 1 for all k and j. (This has a self-reinforcing tendency on the estimates: once all correspondence probabilities equal 0 or 1, there is no internal pressure for $\boldsymbol{\theta}_i^{[n]}$ to change from iteration to iteration.) Furthermore, the $\boldsymbol{\theta}_i^{[n]}$ that take the longest to converge are those for which all or most $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ tend to 0. These hypothesized primitives have little or no empirical evidence for existence, and as such their base amplitude estimates tend to be so low as to make their detection in any image unlikely. Although such primitives would eventually converge and be removed by the model-order adjustment stage, we preemptively remove any primitive whose base amplitude drops below a certain threshold η_a (chosen so that $P_D(\eta_a)$ is near 0). This speeds the ultimate convergence of the algorithm. As a practical matter, we also impose a maximum number of iterations. If this limit is reached, the last iteration's estimate $\boldsymbol{\theta}^{[n]}$ is subjected to one last model-order adjustment step, and the primitives that survive are used as the final estimate.

4.7.3 Numerical Issues

Maximization over location and radius as in (4.28) and (4.29) requires inversion of a matrix formed as a weighted sum of deficient-rank matrices $\mathbf{H}_k^T\mathbf{R}^{-1}\mathbf{H}_k$ or $\tilde{\mathbf{H}}_{i,k}^T\mathbf{R}^{-1}\tilde{\mathbf{H}}_{i,k}$, with weights given by $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$. If $\boldsymbol{\theta}_i^{[n]}$ is a poor estimate or if primitive i is spurious, then most of these correspondence probabilities will be nearly zero, and the inverting matrix may be poorly conditioned. Additionally, even if multiple correspondence probabilities are nonzero, in degenerate cases the sum of these matrices might not achieve full rank. If the condition number [43] of the inverting matrix rises above a certain value at any iteration, the location and radius estimates are not updated at that iteration.[10]

Another numerical issue arises in the generation of the coarse grid of points on the viewsphere for the pose maximization. These grid points are generated according to a deterministic algorithm (the "Hammersley" algorithm of [113]). Observation shows that if the same coarse grid is used repeatedly, estimates of $\theta_i^{\mathbf{P}}$ tend to cluster around the grid points. To correct for this tendency we rotate the coarse-level grid by a random angle before conducting each coarse-to-fine pose search. Similarly, we apply a random offset to the coarse 1-D amplitude grid before performing each coarse-to-fine amplitude search.

[10] In degenerate cases in which the inverting matrix fails to achieve full rank despite multiple nonzero correspondence probabilities, we update the observable-subspace component of the primitive's location and radius estimates.

Chapter 5

Experimental Results from the Basic Algorithm

IN Chapter 3 we developed a framework enabling consideration of target model generation as a data association problem; in Chapter 4 we constructed an algorithm that operates in that framework to generate target models directly from SAR imagery. In this chapter we present experimental results obtained by running the algorithm on several targets. These results illustrate the basic utility of the algorithm and demonstrate how the performance of the algorithm depends on primitive observability. We begin in Section 5.1 by describing the basic experimental setup. In Section 5.2 we present results obtained from algorithm runs on single-primitive targets; these results demonstrate the dependence of algorithm performance on primitive observability and provide benchmarks for the more complex targets that follow. Section 5.3 describes a procedure for correcting a location-estimate bias observed in some of the algorithm results. In Section 5.4 we present results from more sophisticated targets and compare these results to single-primitive benchmarks. Section 5.5 concludes the chapter with an examination of the effects of a mismatch between the primitive dimensions used to construct the algorithm's scattering models and the actual dimensions of observed primitives.

■ 5.1 Experimental Setup

Our target model generation algorithm associates features extracted from SAR images to estimate target primitive characteristics. In the following section we describe the construction of the SAR-image data sets used in our experiments. Implementation of the algorithm also requires specification of the values of various parameters; we describe the selection of these parameters in Section 5.1.2.

■ 5.1.1 Construction of SAR-Image Data Sets

The SAR images used in our experiments were produced using XPatch, a powerful electromagnetic-simulation software package capable of accurately simulating radar returns from a wide variety of targets [41, 69]. XPatch combines "shooting and bouncing rays" (SBR) techniques [4, 29, 66] with PO to predict the backscattered field observed

parameter	value
range resolution	0.3 m
cross-range resolution	0.3 m
range pixel spacing	0.2 m
cross-range pixel spacing	0.2 m
center frequency	9.6 GHz
aperture width	3.0°
sidelobe-suppression frequency weighting	−35-dB Kaiser window (range and cross-range)

Table 5.1. Imaging parameters for XPatch simulations.

from facetized target models, *i.e.*, targets described as unions of multiple small triangular planar regions. XPatch has several independent modules that can each be used for electromagnetic scattering simulation; the two most widely used modules are known as XPatchF and XPatchT [26,27].[1] XPatchF is designed primarily for simulating RCS measurements or range profiles from targets at a single fixed viewing angle, while XPatchT is designed to facilitate the simulation of radar measurements across an aperture. We use XPatchT to produce simulated SAR imagery of the targets considered in our experiments. The parameters used to generate these images are given in Table 5.1. (The values of these parameters were chosen based on their similarity to the imaging parameters of widely available data sets, such as the MSTAR data collections [73, 74].)

Because XPatch produces imagery of targets in the clear, we add synthetic natural clutter to each image.[2] Natural clutter observed in real SAR images varies widely in scale and structure depending on the components of the scene being imaged [103, 104]. We synthesize grassy-terrain-type clutter to add to each SAR image. The small size of the distinct scatterers (*i.e.*, individual blades of grass) in a grassy-terrain environment suggests that this type of clutter can be well-modeled as an additive random process independent between resolution cells [103]. Several authors have suggested the generalized K distribution as a physically relevant and empirically accurate model for the distribution of polarimetric complex amplitudes of clutter pixels [6, 48, 49, 119].[3] We

[1] The trailing initials in these modules' names reflect the fact that XPatchF performs computations in the frequency domain, while XPatchT computations are performed in the time domain.

[2] XPatch does include the capability to augment target imagery with various types of simulated natural clutter using the XPatchES module of the package [25]. However, this module is designed primarily for one-time synthesis of clutter for individual SAR images, and not for repeated syntheses of clutter for a large set of SAR images such as we require to facilitate Monte Carlo analysis.

[3] The K distribution is a unimodal heavy-tailed distribution characterized by three parameters related to its shape, mode, and spread. (For multidimensional measurements—*e.g.*, a complex polarimetric pixel amplitude—the spread parameter is a matrix.) The work of [6,48,49,119] suggests that in some cases the K distribution is a better phenomenological model than other distributions commonly used to model clutter, such as log-normal, Rayleigh, and Weibull [53, 103]. The K distribution takes its name from its dependence on the modified Bessel function of the third kind, which is often denoted $K_{\cdot}(\cdot)$.

thus model clutter as an additive K-distributed random process with statistics given by the "grass" parameters of [80]. This model captures much of the empirical behavior of grassy-terrain clutter in polarimetric systems, including the unequal distribution of clutter among the polarimetric channels and the correlation between HH and VV returns [80, 103]. (The SAR image in Figure 3.2 was generated by XPatch using the parameters of Table 5.1 and corrupted by additive K-distributed grassy-terrain clutter as described above.)

Each experiment in this thesis corresponds to a set of algorithm runs using imagery of a specific target. The SAR-image data set is a fixed size for all runs in an experiment, but each run corresponds to a random choice of viewing angles for the images in the data set and a random realization of clutter in each of those images. To achieve the randomization over viewing angle, we select images for each algorithm run randomly from an XPatch-generated superset of 2736 images for each target. This image superset contains one image for each viewing angle on a 2.5° elevation/azimuth grid on a swath of the viewsphere extending in elevation from 5° to 50° and in azimuth from 0° to 357.5°. We characterize each experiment in terms of its average viewsphere sampling density (AVSD), defined as the average spacing between images in the data set if they were to be spread evenly across the 2.5° viewsphere grid. For instance, a 10° AVSD corresponds to a random sampling of 180 images from the superset, exactly enough to form a 10° grid in azimuth and elevation across the viewsphere band described above if the images were selected in a regularly spaced manner rather than randomly. Similarly, a 20° AVSD corresponds to a sampling of 54 images. Each image selected in any algorithm run is corrupted with K-distributed clutter as described above. A single experiment thus represents the average algorithm performance observed for a data set of given size; this averaging removes variations in performance attributable to the inclusion or exclusion of specific viewing angles in the data set, or to specific clutter realizations.

■ 5.1.2 Specification of Algorithm Parameters

Implementation of the algorithm described in Chapter 4 requires specification of various parameters: the measurement-model parameters of Table 3.2, the primitive dimensions used to construct the scattering responses $S_{\theta_i^t}(\cdot)$ (see Section 3.1 and Appendix A), and additional parameters associated with each algorithm stage. We now discuss how these parameters were chosen.

The measurement-model parameters of Table 3.2 characterize the empirical behavior of the feature-extraction procedure of Section 3.2 in the clutter environment described above. The values we use for these parameters are given in Table 5.2. The false alarm rate γ_{FA} is an empirical value calculated by counting the reports extracted from images containing only clutter. The false-alarm amplitude pdf $p_{FA}(\cdot)$ can be represented as a histogram based on these reports' amplitudes; we implement $p_{FA}(\cdot)$ as a Gaussian approximation to this histogram, as depicted in Figure 5.1. The probability-of-detection function $P_D(\cdot)$ represents the observed detection performance of the feature extractor, obtained by running the feature extractor on imagery containing targets of known

quantity	description	value
γ_{FA}	false alarm rate	0.058 FA/m^2
$p_{FA}(\cdot)$	pdf of false-alarm amplitudes	see Figure 5.1
$P_D(\cdot)$	function mapping scatterer amplitude to probability of detection	see Figure 5.2
$\mathbf{R}^{1/2}$	root-covariance of extracted locations for detections	$\begin{bmatrix} 5.0 \text{ cm} & 0 \\ 0 & 5.0 \text{ cm} \end{bmatrix}$
$\boldsymbol{\rho}$	confusion matrix of extracted types for detections	$\begin{bmatrix} 0.78 & 0.01 & 0.17 & 0.97 \\ 0.22 & 0.99 & 0.83 & 0.03 \end{bmatrix}$
σ_a	standard deviation of extracted amplitudes for detections	5.0 dBsm

Table 5.2. Values of measurement-model parameters and functions. These are the values used for the parameters of Table 3.2. These parameters characterize the feature extraction process of Section 3.2 in the clutter environment described in Section 5.1.1.

amplitude for RCS values between -10 dBsm and 20 dBsm at 1-dBsm increments. The resulting $P_D(\cdot)$ function is depicted in Figure 5.2. The extracted-location covariance matrix $\mathbf{R}$ and extracted-type confusion matrix $\boldsymbol{\rho}$ are empirical statistics obtained by running the feature extractor on imagery of 10-dBsm-base-amplitude primitives from various viewing angles. (Note that the trihedral and dihedral both exhibit significant confusions between type classifications, whereas the tophat and cylinder do not. This is due to the relatively significant contributions of the trihedral's and dihedral's secondary-response mechanisms to those primitives' overall responses—*e.g.*, the presence of strong double-bounce reflections in the trihedral response.) The extracted-amplitude variance σ_a^2 was not directly calculated as an empirical statistic as were $\mathbf{R}$ and $\boldsymbol{\rho}$, but is instead a heuristic value chosen with the intent of capturing some of the variability in primitive responses encountered in the real world that would be difficult to model in a training set. We examine more principled and directed approaches to dealing with primitive-response noncanonicity in Chapter 7.

The values in Table 5.2 can be viewed as approximate or average statistics; the "true" values of many of these parameters will vary depending on the characteristics of the targets being imaged. For instance, bright scatterers will produce reports with a high signal-to-clutter ratio, which will tend to improve the quality of the component measurements of the report. Thus we would expect the true covariance of $\mathbf{x}_{k,j}$ and the true confusion of $t_{k,j}$ to vary depending on $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$. Similarly, a primitive's probability of detection in any image might depend not only on the magnitude of its polarimetric response, but also on scatterer type because different polarimetric signatures of identical magnitude might not be equally discernible in clutter that is unequally distributed across polarimetric channels. Additionally, type confusion can depend on primitive dimensions: as a trihedral's dimensions increase, for instance, its double-bounce response mechanisms become more specular and contribute less to the

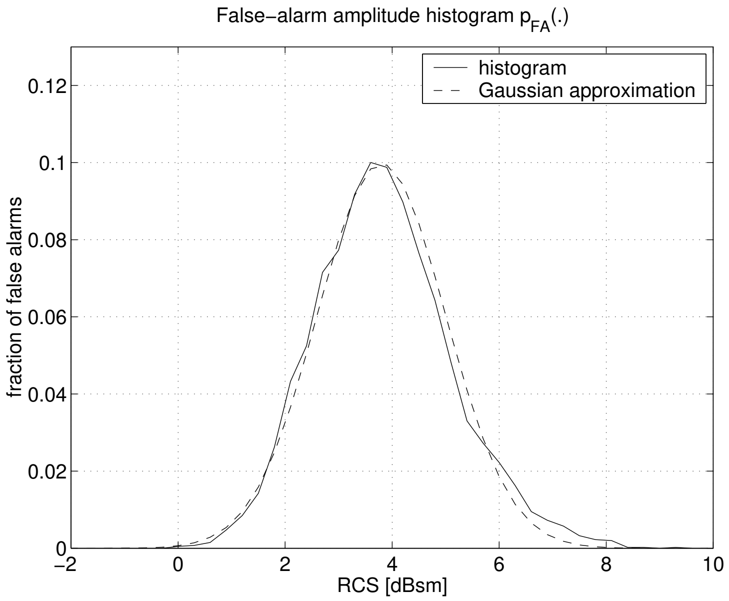


Figure 5.1. False-alarm amplitude pdf $p_{FA}(\cdot)$. The solid line is a histogram of the observed amplitudes of 4000 false alarms. It is implemented as a Gaussian with mean 3.8 and standard deviation 1.2, overlaid as the dashed line.

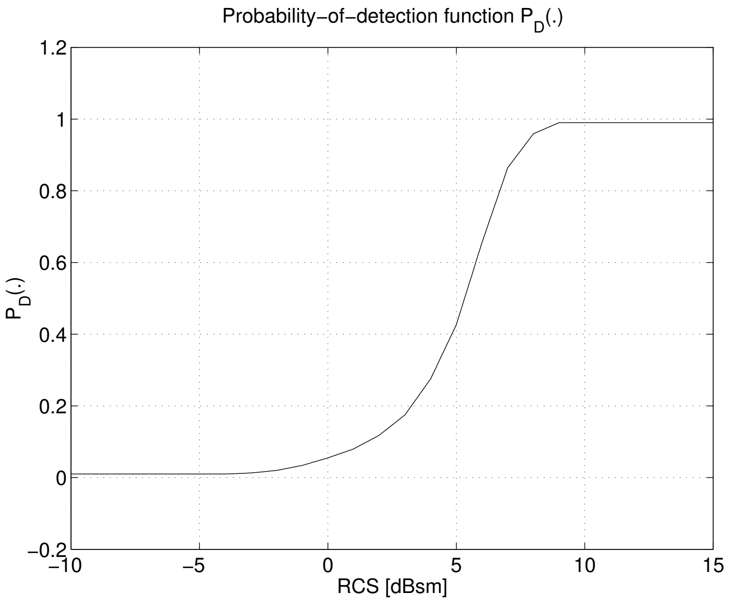


Figure 5.2. Probability-of-detection function $P_D(\cdot)$. This function was empirically estimated by running the feature extractor on imagery containing targets of known amplitude at 1-dBsm increments. It is linearly interpolated between those points. For robustness it is clipped at a maximum value of 0.99 and a minimum value of 0.01 regardless of amplitude.

overall scattering response. It is possible to extend the measurement model developed in Section 3.3 to capture these dependences; such an extension is discussed further in Chapter 8.

Implementation of the model generation algorithm also requires specification of nominal primitive dimensions to construct scattering models $S_{\theta_i^t}(.)$ as described in Section 3.1 and Appendix A. For most of the experiments in this section, the nominal scattering-model primitive dimensions have been set to values that are identical to the dimensions of the primitives to be estimated. These values are indicated in the context of each experiment. In Section 5.5 we examine the effects of a mismatch between modeled and actual primitive dimensions. These effects are seen to be relatively minor.

Finally, we must also specify various algorithmic parameters that are required for implementation of the different stages of the algorithm. These parameters, their notation, and their values are given in Table 5.3. As indicated in the various sections of Chapter 4, these parameters are heuristically chosen to give reliable performance. The selection of these parameters is summarized below.

- *Initialization stage.* As described in Section 4.4, implementation of the initialization stage requires specification of several threshold parameters: an amplitude threshold $a_{\min}$ for deciding which reports to pass to the clusterer, a chi-squared-agglomeration dissimilarity threshold η indicating how dissimilar reports can be and still be grouped, and a report-count threshold $n_{\min}$ used to weed out spurious clusters at the conclusion of the agglomeration. Selection of $a_{\min}$ is essentially the choice of a point on a receiver-operating characteristic (ROC) curve [105] indicating the relative abundance of missed detections and false alarms among the reports passed to the clusterer. We choose $a_{\min}$ to be 5.0 dBsm, because most false alarms have amplitudes below this value (see Figure 5.1) and because primitives much dimmer than this are unlikely to produce a report (see Figure 5.2). The chi-squared-statistic dissimilarity threshold η is set to 0.9, a value empirically determined to be large enough to group most reports attributable to a single primitive but not so large as to group many unrelated reports.[4] The report-count threshold $n_{\min}$ is chosen based on the AVSD, since a sparser AVSD (*i.e.*, a smaller K) will generally result in fewer reports attributable to a given primitive.

- *E step.* The gate radius is chosen to be 0.4 m, which is sufficiently large to accommodate not only the covariance of $\mathbf{x}_{k,j}$ (see $\mathbf{R}$ in Table 5.2) but also errors in the location estimates $\theta_i^{\mathbf{x}}$ observed early in the EM iteration.

- *Model-order adjustment stage.* As described in Section 4.5, the type-dependent probability-of-detection thresholds are selected to reflect the fact that tophats

[4]Choosing a relatively large value for η aids in properly clustering reports generated by a single primitive but emanating from points in $\mathbb{R}^3$ other than $\theta_i^{\mathbf{x}}$, *e.g.*, reports from the wandering specular-reflection points of cylinders and tophats, and lower-bounce reports from trihedrals and dihedrals. (The inability of the agglomerative clustering procedure to model wandering-specular-point reflections properly was discussed in Section 4.4).

algorithm stage	parameter	notation	value
initialization	minimum report amplitude for clustering	$a_{\min}$	5.0 dBsm
	chi-squared-statistic dissimilarity threshold	η	0.9
	minimum reports in cluster for initialization	$n_{\min}$	$\begin{cases} 4, & \text{AVSD} = 10^\circ, \\ 3, & \text{AVSD} = 20^\circ \end{cases}$
E step	gate radius	r_{gate}	0.4 m
model-order adjustment	empirical-probability-of-detecton thresholds	$\tilde{p}^{d}_{\min}(\theta_i^{t\,[n]})$	$\begin{cases} 0.06, & \theta_i^{t\,[n]} = 1, \\ 0.2, & \theta_i^{t\,[n]} = 2, \\ 2.9/K, & \theta_i^{t\,[n]} = 3, \\ 2.9/K, & \theta_i^{t\,[n]} = 4 \end{cases}$
termination check	location estimate tolerance	$\tau_{\mathbf{x}}$	1.0 cm
	radius estimate tolerance	τ_r	1.0 cm
	pose az/el estimate tolerance	$\tau_{\mathbf{p}_1}$	2°
	pose rotation estimate tolerance	$\tau_{\mathbf{p}_2}$	2°
	base amplitude tolerance	τ_a	0.3 dBsm

Table 5.3. Values of algorithmic parameters.

and trihedrals have much broader responses than dihedrals and cylinders, and will usually produce more reports in a given data set. The thresholds for dihedrals and cylinders depend on K (the number of images in the data set) so that regardless of the size of the data set, primitive estimates based on as few as three reports can survive the model-order adjustment stage.[5]

- *Termination check.* Convergence is determined by comparing estimate values between successive iterations as described in Section 4.6. The tolerances on the continuous parameters of $\boldsymbol{\theta}_i$ are chosen to be small values as indicated in Table 5.3.

■ 5.2 Single-Primitive Target Experiments

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. The purpose of these experiments is threefold: to validate algorithm performance on a set of very simple targets, to illustrate the limitations in performance imposed by the relative differences in observability between the four primitive types, and to tabulate benchmarks against which results from more complex targets can be judged. Although many primitives observed in practice will have base amplitudes much greater than 10 dBsm (especially dihedrals and trihedrals in many targets of interest [62]), direct comparison of the

[5] At least two $\mathbf{x}_{k,j}$ are required to form an estimate of a dihedral's location; at least three $\mathbf{x}_{k,j}$ are required to form estimates of a cylinder's location and radius.

primitive	dimensions
trihedral	$a = 12.67$ cm
tophat	$r = 18.39$ cm, $h = 36.78$ cm
dihedral	$a = 14.05$ cm, $b = 14.05$ cm
cylinder	$r = 17.70$ cm, $h = 53.04$ cm

Table 5.4. Dimensions of single-primitive targets. (The dimension notation used here corresponds to that of Figure 2.10.) These dimensions give a base amplitude of 10 dBsm for each primitive.

target	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}T}$ [cm]	θ_i^r [cm]	$\theta_i^{\mathbf{p}T}$ [°]
trihedral	1	10	[30.5 0.0 15.2]	—	[35.26 0 0]
tophat	2	10	[30.5 0.0 15.2]	18.39	[90 0 —]
dihedral	3	10	[30.5 0.0 15.2]	—	[25 0 0]
cylinder	4	10	[30.5 0.0 15.2]	17.70	[65 180 —]

Table 5.5. Parameters of single-primitive targets.

four primitive types with the same base amplitude helps to illustrate the differences in observability between the primitives. The dimensions of the primitives used in the experiments of this section are given in Table 5.4. As indicated previously, the $S_{\theta_i^t}(\cdot)$ in these experiments are also constructed using the dimensions in Table 5.4; the effects of dimension mismatch are examined in Section 5.5.

The components of $\boldsymbol{\theta}_i$ for each of the four single-primitive targets are listed in Table 5.5. Recall that specification of the pose of a tophat or cylinder requires two Euler angles, while specification of the pose of a trihedral or dihedral requires three Euler angles (see Section 3.1). The two components of the tophat and cylinder pose correspond to the elevation and azimuth of these primitives' axis of symmetry. The first two components of the trihedral and dihedral pose correspond to the elevation and azimuth of the direction of maximum RCS (*i.e.*, the viewing angle at which $S_{\theta_i^t}(\cdot)$ is maximized); the third component is the rotation of the response about this direction. The trihedral and tophat are thus oriented with their bases parallel to the ground plane, while the dihedral and cylinder are tilted 25° out of vertical.

Before presenting results for the algorithm runs on these targets, we point out that the differences in the relative observabilities of these primitives will lead to marked differences in the performance of the algorithm on each. Figure 5.3 depicts the predicted RCS values $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ of each of the four primitives described by the parameters of Table 5.5 at the 2736 viewing angles in the 2.5° grid. This figure clearly depicts the differences in primitive observability. The tophat has the broadest response and would produce a report from almost any viewing angle. The trihedral response is localized to roughly a quarter of the grid but is still more observable than the dihedral and cylinder, both of which are observable only near a single great circle along the viewsphere corresponding to each primitive's specular plane. Clearly, the narrow responses of the dihedral and cylinder indicate that these primitives will generally produce many fewer reports in a given data set than will tophats or trihedrals. As a result the algorithm will be more likely to discover the presence of a tophat or trihedral in a data set of given size than to discover the presence of a dihedral or cylinder.[6] Additionally, dihedrals and cylinders will generally suffer from poorer-quality estimates of $\boldsymbol{\theta}_i$ (compared to those of tophats and trihedrals) even when they are discovered.

We can quantify the relative observabilities of these primitives by computing the average $P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))$ over the 2.5° grid for each primitive. This gives the fraction of images in a data set in which each primitive can be expected to produce a report. Denoting this quantity as $\overline{P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))}$, we see that $K \cdot \overline{P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))}$ is the average number of reports a primitive can be expected to produce in a randomly selected data set of size K. Table 5.6 lists these values for 10° and 20° AVSDs. We can also apply $P_D(\cdot)$ to the $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ grids to obtain a rough upper bound $\tilde{P}_{\text{disc}}^{\max}$ on the fraction of runs in any experiment in which it is possible for the algorithm to discover the presence of the primitive. (Recall from Section 4.4 that discovery of a primitive requires that the data set contain at least $n_{\min}$ reports of amplitude $a_{\min}$ or greater.) We obtain $\tilde{P}_{\text{disc}}^{\max}$ for each data-set size by Monte Carlo simulation: we randomly select K viewing angles from the 2.5° grid and then randomly "detect" or "miss" the primitive at each sample based on $P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))$ at that viewing angle. If at least $n_{\min}$ "detections" are produced at viewing angles for which $A(\boldsymbol{\theta}_i, \psi_k, \phi_k) > a_{\min}$, it would be possible for the algorithm to discover the primitive. The fraction of simulations in which this is the case provides the approximate bound $\tilde{P}_{\text{disc}}^{\max}$.[7] (Note that discovery performance near $\tilde{P}_{\text{disc}}^{\max}$ would indicate that the initialization stage is performing as intended.) Table 5.6 indicates that the tophat is by far the most observable primitive: it will produce many reports and should be discoverable in almost all algorithm runs. The trihedral will also generally produce enough reports to be discovered in most trials. The dihedral is much less observable than the tophat or trihedral, but should still be discovered in more algorithm runs than the cylinder; both the dihedral and cylinder, however, have a P_{disc} bound that is significantly below unity and deteriorates with AVSD.

Table 5.7 presents the performance of the algorithm on the four single-primitive targets for 10° and 20° AVSDs. (We analyze this performance in detail below.) This table is broken up into several sections, each describing the errors observed in a single component of the estimate of $\boldsymbol{\theta}$. Model order results are broken into three categories:

[6] We use the term "discovery" rather than "detection" to refer to the production of a primitive estimate by the algorithm in order to avoid confusion with the detection of reports in individual images. Primitive discovery requires detection in multiple images, but detections will not of necessity lead to discovery.

[7] The bound obtained in this way is approximate because it neglects effects such as clutter and mismatch between $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ and XPatch predictions (described in conjunction with the algorithm's base amplitude estimate performance shortly). Additionally, it neglects the presence of false alarms, which could be grouped with $(n_{\min} - 1)$ legitimate reports to produce a cluster that survives the initialization stage.

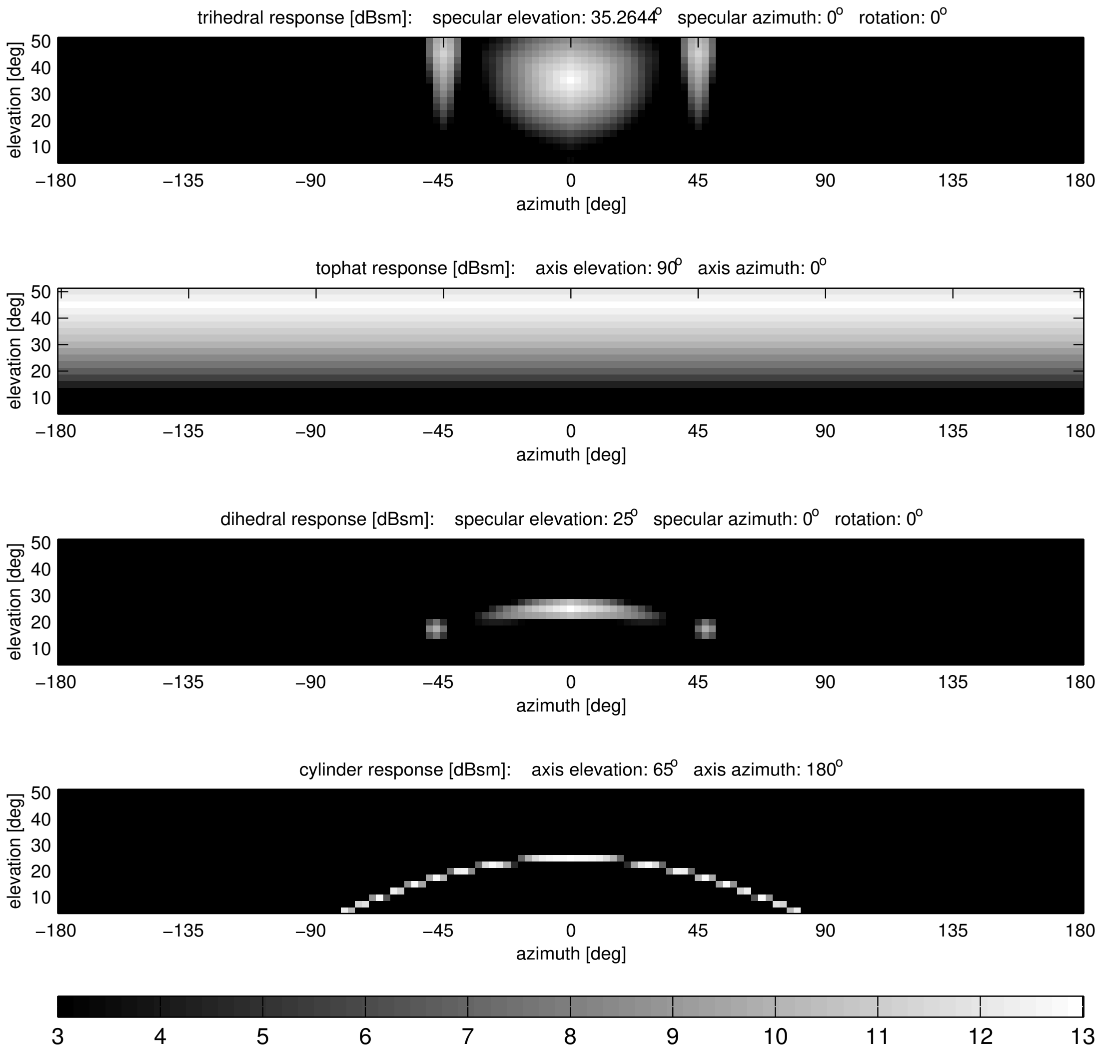


Figure 5.3. Responses of four primitives. These images depict each primitive's $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ at every point on the 2.5° viewsphere grid. (The images are clipped below 3 dBsm because primitives are not detectable from viewing angles at which $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ is too low.) The relative observabilities of the different primitive types, as clearly visible above, have a significant impact on the performance of the algorithm.

primitive	$\overline{P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))}$	$K \cdot \overline{P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))}$		$\tilde{P}_{\text{disc}}^{\max}$	
		10° AVSD	20° AVSD	10° AVSD	20° AVSD
trihedral	0.1349	24.28	7.29	1.000	0.962
tophat	0.7287	131.17	39.35	1.000	1.000
dihedral	0.0292	5.26	1.58	0.683	0.154
cylinder	0.0237	4.26	1.28	0.585	0.129

Table 5.6. Relative observability of primitive responses. The quantity $\overline{P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))}$ is calculated from the primitive responses of Figure 5.3 and is used to determine the average number of reports that will be produced by each primitive in 10°- and 20°-AVSD data sets. The rightmost two columns present rough upper bounds on P_{disc}, the fraction of algorithm runs in which a primitive is discovered. These values suggest that while tophats and trihedrals will be discoverable in almost all trials, dihedrals and cylinders will often not produce enough reports to be discovered.

"underfit" refers to the the fraction of runs in which the model order was zero, *i.e.*, in which the presence of the primitive was not discovered by the algorithm; "P_{disc}" 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 adjustment 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 discovery. In most algorithm runs in which the primitive is discovered 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. The errors in the estimates of the continuous parameters of $\boldsymbol{\theta}_i$, conditional on discovery, correct type identification, and non-spurious parameters, are presented in the lower half of Table 5.7. Base amplitude figures are quoted in dBsm, location and radius in centimeters, and pose in degrees. We separate location and base amplitude statistics into bias (defined here as average estimate minus truth) and standard deviation to illustrate bias effects described below. Pose results are presented in terms of two statistics, the first being the root-mean-squared error (rmse) of the azimuth/elevation component (*i.e.*, the root-mean-squared angular separation on the viewsphere between the azimuth/elevation estimate and the true azimuth/elevation) and the second being the rmse of the rotation component, if applicable. 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 discovered and estimated to be the correct type.

Discoverability and Model-Order Estimation

The P_{disc} statistics of Table 5.7 confirm that algorithm performance depends strongly on primitive observability. These statistics also indicate that the algorithm is discovering each primitive in almost all cases in which it is provided enough reports to do

AVSD	target	model order			type confusion	fraction spurious
		underfit	P_{disc}	overfit		
10°	trihedral	0.000	1.000	0.000	[1.000 0.000 0.000 0.000]	0.000
	tophat	0.000	1.000	0.000	[0.000 1.000 0.000 0.000]	0.000
	dihedral	0.279	0.721	0.000	[0.000 0.000 0.998 0.002]	0.002
	cylinder	0.051	0.490	0.000	[0.000 0.000 0.021 0.979]	0.002
20°	trihedral	0.052	0.948	0.000	[0.990 0.000 0.006 0.004]	0.000
	tophat	0.000	1.000	0.000	[0.000 1.000 0.000 0.000]	0.000
	dihedral	0.815	0.185	0.000	[0.035 0.000 0.933 0.032]	0.000
	cylinder	0.891	0.109	0.002	[0.030 0.000 0.004 0.966]	0.026

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

Table 5.7. Results from single-primitive-target experiments.

so: the trihedral, tophat, and dihedral P_{disc} results are all very close to the bounds of Table 5.6. (The observed P_{disc} actually exceeds the approximate bound for both dihedral experiments; this is attributable to the slight mismatch between the dihedral primitive scattering model and the PO predictions as described in Section 3.1 and Appendix A.) The cylinder P_{disc} values are slightly below the bound for both AVSDs. This is attributable to the inability of our initialization procedure to deal directly with the cylinder's wandering specular-reflection point as noted at the conclusion of Section 4.4. In particular, in some algorithm runs in which a cylinder does produce n_{min} or more reports of amplitude a_{min}, those reports are not properly clustered by the initialization stage.[8]

We see from Table 5.7 that model order is almost never overestimated for single-primitive targets, the exception being a single run in the 20° AVSD cylinder experiment. In most cases in which the primitive is discovered, the initialization stage produces a single initialization that persists to the end of the EM iteration. In cases where the initialization stage does produce more than one report cluster, the spurious cluster is almost always removed in a subsequent iteration by the model-order adjustment stage. Additionally, the fact that the algorithm is approximately meeting the P_{disc} bounds of Table 5.6 indicates that an underfit in model order is almost always attributable to a paucity of reports and not to reports being incorrectly associated. The bottom line is that the model-order adjustment stage is properly removing most spurious estimates but not at the expense of removing primitives that correspond to true primitives.

Type Estimates

It is apparent from Table 5.7 that the type classification performance of the algorithm is excellent: in almost every trial in which a primitive is discovered, its type is correctly identified. This suggests that the limited type information provided by the even-bounce/odd-bounce discriminator of the feature extraction stage is not a significant impediment to type estimation. In the 10° AVSD experiments, there is very limited confusion between the dihedral and cylinder responses; these two responses are both concentrated near a single plane and are sufficiently similar (see Figure 5.3) that a small set of measurements might not be sufficient to distinguish them reliably. In some of the the 20° AVSD experiments, dihedrals and cylinders were also incorrectly identified as trihedrals. Note that in no case was any primitive incorrectly identified as a tophat: the tophat response is so broad that it is a poor fit to any other primitive response even at a 20° AVSD.

Base Amplitude Estimates

The observed standard deviations of the base amplitude estimates (as well as those of the location and radius estimates) indicate that, as expected, the more-observable prim-

[8]Although the tophat also has a wandering wandering specular-reflection point, its broad response ensures that at least n_{min} reports will be grouped even if some reports are not properly clustered by the initialization.

primitive	average XPatch $-$ $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$
trihedral	-0.30 dBsm
tophat	-1.35 dBsm
dihedral	-0.33 dBsm
cylinder	-0.49 dBsm

Table 5.8. Average differences between XPatch predictions and $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$. These averages are over all viewing angles in which the maximum XPatch-predicted pixel amplitude exceeds 3 dBsm.

itives enable more accurate parameter estimation because they usually produce more reports in any data set. Table 5.7 also indicates that the base amplitude estimates are generally biased. This is attributable to several factors. First is the frequency windowing inherent in the SAR imaging process. As indicated in Section 2.2, a primitive's amplitude in an image is affected by its location in the slant plane relative to the pixel centers: a primitive whose projected location is not coincident with a pixel center will be slightly attenuated, *i.e.*, the extracted amplitude $a_{k,j}$ will be less than $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$. The average attenuation will depend on the oversampling rate and on the sidelobe-suppression windowing function. For the parameters of Table 5.1 this average attenuation can be shown to be approximately 0.5 dBsm.[9] Another factor influencing the base amplitude bias is the slight deviation between $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ and the RCS values produced by XPatch. These differences arise from several sources, including approximations made in our construction of $S_{\theta_i^t}(\cdot)$ from basic scattering-response components (see Section 3.1 and Appendix A), slight inaccuracies in the basic scattering-response components themselves[10] [62,92], and SBR and facetization effects [4,27,66]. The mean difference between the XPatch prediction and $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ over the 2.5° grid for each primitive type is given in Table 5.8. (Note that the presence of the nonlinear $P_{D'_{k,i}}$ terms in (4.24) indicates that the biases of Table 5.8 will not necessarily be the same as the biases observed in the base amplitude estimates.) A final factor influencing the base amplitude bias is the correlation of the base amplitude estimate error with pose estimate error due to these parameters' coupling in each $S_{\theta_i^t}(\cdot)$. This correlation is most pronounced for the cylinder due to its extremely narrow response: if the pose estimate converges to a slightly erroneous value, the estimate of θ_i^a will tend to increase in order to better fit the $a_{k,j}$ measurements, all of which lie near the true specular plane. As long as this increase is small enough not to boost the $P_{D'_{k,i}}$ values at other viewing angles, there will be no pressure in (4.24) to counteract this tendency.

[9]This value is obtained by numerical integration of the -35-dB Kaiser sidelobe-suppression kernel over a square region determined by the oversampling rate.

[10]For instance, the shaping function of the tophat double-bounce in Table 2.1 depends on an approximation to the projected area of the tophat double-bounce region, which is complicated due to the curved contours of the base plate and cylinder. Exact determination of this area at any viewing angle requires numerical integration [12].

Location and Radius-of-Curvature Estimates

As expected, the location and radius-of-curvature estimates produced by the algorithm are most accurate in high-observability situations, *i.e.*, for denser AVSDs and for primitives with more-observable responses. Note that due to the high specularity of the dihedral and cylinder responses, all reports produced by either primitive will be nearly coplanar and will thus provide little information about the component of $\theta_i^{\mathbf{x}}$ orthogonal to the specular plane. Dihedral and cylinder location-estimate standard deviations are thus disproportionately large, even for their low observability. Additionally, although it is not illustrated in Table 5.7, the radius errors are correlated with the location errors due to layover effects.

The trihedral and dihedral location estimates exhibit a pronounced bias. This is due to the influence of these primitives' lower-bounce mechanisms as described in Section 3.1. In particular, these lower-bounce reflections violate the model of (3.3) that is the basis for the estimation of $\theta_i^{\mathbf{x}}$ in (4.28). In Section 5.3 we discuss an approach to removing this bias, if desired.

Pose Estimates

The pose results of Table 5.7 demonstrate that primitive pose can in general be accurately estimated to a finer granularity than provided by the AVSD. As with the other parameters, pose estimation improves with observability; however, the distinctive nature of each primitive response (see Figure 5.3) has different implications for the estimation of each primitive's pose. In particular, tophats produce many reports around the viewsphere and possess rotational symmetry, making their responses easy to orient. Cylinders produce very few reports, but all of these reports are nearly coplanar, making cylinder pose estimation relatively easy as well.

The trihedral and dihedral responses suffer from specific invariances that complicate the estimation of their poses. For instance, estimation of the azimuth and elevation of the trihedral is relatively easy, but estimation of its rotation is relatively difficult because the trihedral response is nearly invariant to rotation about its specular direction.[11] The dihedral suffers from a large pose error due to its confinement near a single specular plane: it is difficult to distinguish dihedral poses related by a small rotation within this plane. Figure 5.4 is a visual depiction of this phenomenon. This figure presents four grids of $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ measurements, each obtained from a dihedral identical to the one under consideration but oriented at a different pose. The poses of the dihedrals in successive images in Figure 5.4 are related by rotations of approximately 5° within the specular plane. Visually, these responses are very similar; it is clear that a sparse subsampling of the images in Figure 5.4 might provide insufficient information for reliable discrimination between these responses.

[11] This problem is exacerbated for larger trihedrals, which have narrower double-bounce components and thus provide less rotational information.

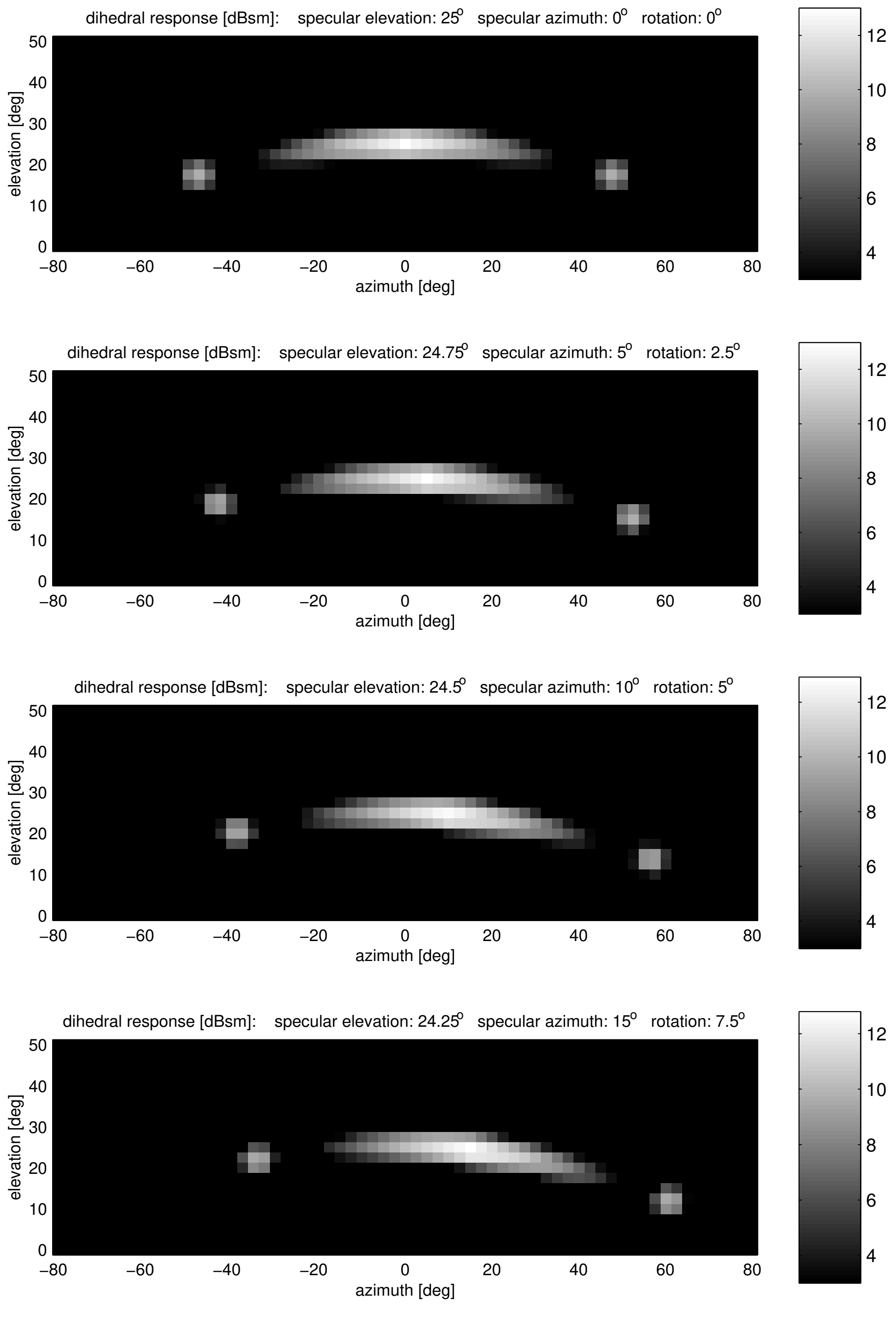


Figure 5.4. Responses of four identical dihedrals with different poses. (The format of the images in this figure is similar to that of Figure 5.3.) Orientations of the dihedrals in successive images are related by a rotation of approximately 5° within the specular plane. This illustrates a near-invariance in dihedral pose: changes in pose can produce responses that are difficult to distinguish, especially given measurements over only a sparse sampling of the above grids.

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

Table 5.9. Results of bias-reduction post-processing on trihedral and dihedral location estimates.

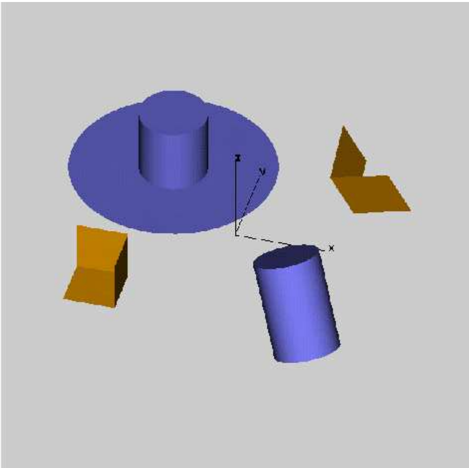
5.3 Reducing Trihedral and Dihedral Location Estimate Bias

As described in the previous section, the location estimates of the trihedral and dihedral, formed according to (4.28) at each iteration, are biased due to the influence of lower-bounce reflections. The direction and magnitude of the bias depends on the primitive orientation and dimensions. If we seek an unbiased location estimate, we could modify the projectional model of (3.3) to take into account the different reflection points of the lower-bounce mechanisms. Unfortunately, this conceptually optimal approach introduces a dependence on primitive dimension and pose and significantly complicates the maximization at each iteration. We opt for a computationally simpler approach that is in effect a post-processing step to be implemented after the EM iteration has converged. Specifically, at the conclusion of the EM 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 M-step pose to identify reports obtained at viewing angles from which the highest-bounce response is believed to dominate; we then use the E-step probabilities and $\mathbf{x}_{k,j}$ measurements of these reports to form a new estimate as in (4.28). In other words, we use the final pose estimate to identify reports whose slant-plane location measurements are believed to violate (3.3) and then form a new estimate of $\theta_i^{\mathbf{x}}$ omitting those reports. If the pose estimate is accurate, this refined location estimate will be unbiased (though by omitting some $\mathbf{x}_{k,j}$ we will likely increase the rmse of the location estimate).

The results of this post-processing on the experiments of the previous section are presented in Table 5.9. In each case the bias in the location estimate is reduced significantly. The trihedral location-estimate 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 5.7). For similar reasons, the refinement is more successful for the $10°$ AVSD experiments than for the $20°$ AVSD experiments.

5.4 Multiple-Primitive Target Experiments

The results of Section 5.2 demonstrate the performance of the algorithm on four single-primitive targets, for which the underlying data association problem is quite simple. In this section we examine the performance of the algorithm when it is faced with



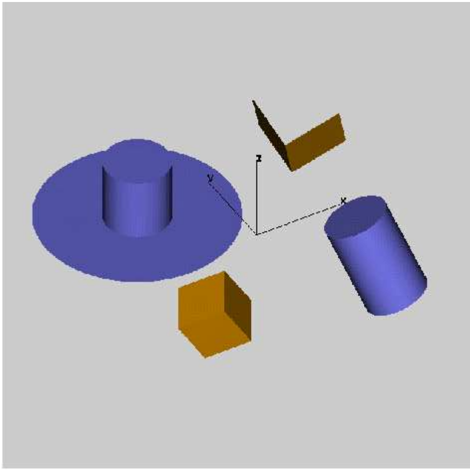


Figure 5.5. Two views of the first multiple-primitive target.

a more challenging data association problem—specifically, a multiple-primitive target. We would expect the error statistics obtained for each primitive in a multiple-primitive target to be no better than those obtained from a target containing only that primitive, *i.e.*, in the absence of confusing primitives. Similar error statistics would indicate that the algorithm is successfully solving the data association problem, correctly assigning report correspondences to arrive at its parameter estimates. In this section we compare results obtained from two multiple-primitive targets with benchmark results obtained from single-primitive targets.

■ 5.4.1 Target One

The first multiple-primitive target we examine consists of four primitives, one of each type, centered at the corners of a 0.91 m square. Each primitive is thus separated from its neighbors by at most three resolution cells.[12] Two renderings of the XPatch facetization model of this target are depicted in Figure 5.5. The cylinder and tophat in this scene are identical to those of Section 5.2; the dihedral and trihedral are larger than those of Section 5.2. (The dimensions of these primitives are given in Table 5.10.) The parameterizations of the primitives in this target are given in Table 5.11. As in Section 5.2, the experiments of this section are conducted using $S_{\theta_i^t}(\cdot)$ matched to the primitive dimensions, and examination of the effects of mismatch is deferred to Section 5.5.

The single-primitive-target benchmark results for the components of this target are given in Table 5.12. Because the tophat and cylinder in this target are the same size as those in Section 5.2, we can use the tophat and cylinder results from that section as benchmarks; to obtain benchmarks for the larger trihedral and dihedral used here, we

[12]From most viewing angles, neighboring primitives will be separated by less than 0.91 m in the slant plane due to the projection from $\mathbb{R}^3$.

primitive	dimensions
trihedral	$a = 28.00$ cm
tophat	$r = 18.39$ cm, $h = 36.78$ cm
dihedral	$a = 33.30$ cm, $b = 33.30$ cm
cylinder	$r = 17.70$ cm, $h = 53.04$ cm

Table 5.10. Dimensions of the components of the first multiple-primitive target. (The dimension notation used here corresponds to that of Figure 2.10.) These dimensions give a base amplitude of 10 dBsm for the tophat and cylinder, 23.75 dBsm for the trihedral, and 25 dBsm for the dihedral.

target component	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}T}$ [cm]	θ_i^r [cm]	$\theta_i^{\mathbf{p}T}$ [°]
trihedral	1	23.75	[−45.7 −45.7 0]	—	[35.26 135 0]
tophat	2	10.00	[−45.7 45.7 0]	18.39	[90 0 —]
dihedral	3	25.00	[45.7 45.7 0]	—	[25 −45 0]
cylinder	4	10.00	[45.7 −45.7 0]	17.70	[65 −135 —]

Table 5.11. Parameters of the components of target one.

ran experiments similar to those of Section 5.2 for these two primitives.

We performed 500 Monte Carlo runs for both a 10° AVSD and a 20° AVSD for this target.[13] The results from these experiments are presented in Table 5.13 in a format similar to that of Table 5.12. (For a multiple-primitive target, overfit statistics cannot be tabulated for individual components. Spurious-estimate statistics are also listed for the target as a whole, and do not include those runs which produced an overfit.) Comparison to the benchmark results of Table 5.12 shows that the error statistics for each component of this target are almost identical to the corresponding single-primitive benchmarks. In particular, the type confusion results are almost identical, all base amplitude biases and standard deviations are within a fraction of a dBsm from their benchmark values, and the location, pose, and radius errors are all very similar to the benchmarks. (The greatest deviation from the benchmark values is observed in the cylinder results from the 20° AVSD performance. This is largely attributable to the lesser statistical significance of the cylinder results in Table 5.13: these statistics are based on a total of $500 \times 0.082 = 41$ cylinder discoveries, compared to 500 discoveries in the single-primitive-target benchmark.) Additionally, each primitive's P_{disc} is almost as high in this experiment as in the single-primitive-target benchmarks. The exception is the cylinder in the 10° experiment, whose P_{disc} decreases from 0.490 to 0.406. This suggests that in approximately 8.4% of algorithm runs on this target, one or more cylinder reports are improperly clustered. Because of the low observability of

[13]Note that this differs slightly from the single-primitive-target experiments, in which Monte Carlo runs were continued until the primitive had been discovered and correctly identified 500 times.

AVSD	target component	model order			type confusion	fraction spurious
		underfit	P_{disc}	overfit		
10°	trihedral	0.000	1.000	0.002	[1.000 0.000 0.000 0.000]	0.000
	tophat	0.000	1.000	0.000	[0.000 1.000 0.000 0.000]	0.000
	dihedral	0.047	0.953	0.000	[0.002 0.000 0.996 0.002]	0.000
	cylinder	0.051	0.490	0.000	[0.000 0.000 0.021 0.979]	0.002
20°	trihedral	0.000	1.000	0.010	[1.000 0.000 0.000 0.000]	0.000
	tophat	0.000	1.000	0.000	[0.000 1.000 0.000 0.000]	0.000
	dihedral	0.563	0.437	0.000	[0.023 0.002 0.961 0.014]	0.000
	cylinder	0.891	0.109	0.002	[0.030 0.000 0.004 0.966]	0.026

AVSD	target component	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse [cm]
		bias	stdev	$\lVert\text{bias}\rVert$	$\sqrt{\text{tr(cov)}}$	az/el	rot	
10°	trihedral	−0.703	0.541	6.799	2.211	2.890	19.246	—
	tophat	−1.382	0.191	0.409	1.288	1.329	—	1.056
	dihedral	−2.033	2.338	5.764	13.890	11.713	9.281	—
	cylinder	0.080	1.625	0.410	22.169	1.286	—	6.629
20°	trihedral	−0.865	0.914	6.271	4.744	4.641	25.058	—
	tophat	−1.298	0.365	0.446	2.175	2.388	—	1.730
	dihedral	−3.225	4.423	4.862	23.908	15.660	14.963	—
	cylinder	0.455	2.771	1.786	30.845	3.433	—	9.980

Table 5.12. Single-primitive-target benchmarks for comparison to results from target one.

AVSD	target component	P_{disc}	type confusion	overfit	fraction spurious
10°	trihedral	1.000	[1.000 0.000 0.000 0.000]	0.020	0.008
	tophat	1.000	[0.000 1.000 0.000 0.000]		
	dihedral	0.950	[0.000 0.000 0.998 0.002]		
	cylinder	0.406	[0.000 0.000 0.045 0.955]		
20°	trihedral	1.000	[1.000 0.000 0.000 0.000]	0.000	0.016
	tophat	1.000	[0.000 1.000 0.000 0.000]		
	dihedral	0.460	[0.026 0.004 0.953 0.017]		
	cylinder	0.082	[0.024 0.000 0.024 0.952]		

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

Table 5.13. Results from first multiple-primitive-target experiments.

the cylinder (see Figure 5.3 and Table 5.6), a single misclustered report could easily be the difference between a cluster that is used to initialize a primitive estimate and a cluster that is removed because it does not contain enough reports.

The similarity between the multiple-primitive-target experiment statistics and the single-primitive-target benchmarks suggests that the algorithm is successfully performing data association. The initialization clustering and subsequent EM iteration are thus finding the proper correspondences between reports in different images and using these correspondences to estimate primitive parameters.

■ 5.4.2 Target Two

The second multiple-primitive target we examine also consists of four primitives. The components of this target are all trihedrals, each identical in size to the trihedral component of the target in the previous section. Two renderings of the XPatch facetization model of this target are depicted in Figure 5.6. The components of this target are located in closer proximity than those of target one: two of the trihedrals are separated by only 30.5 cm, or approximately the width of one resolution cell. The parameterizations of the components of this target are given in Table 5.11. The primitive dimensions used to construct the $S_{\theta_i^t}(\cdot)$ for this experiment are those of Table 5.10.

We performed 125 Monte Carlo runs at 10° and 20° AVSDs for this target. Because

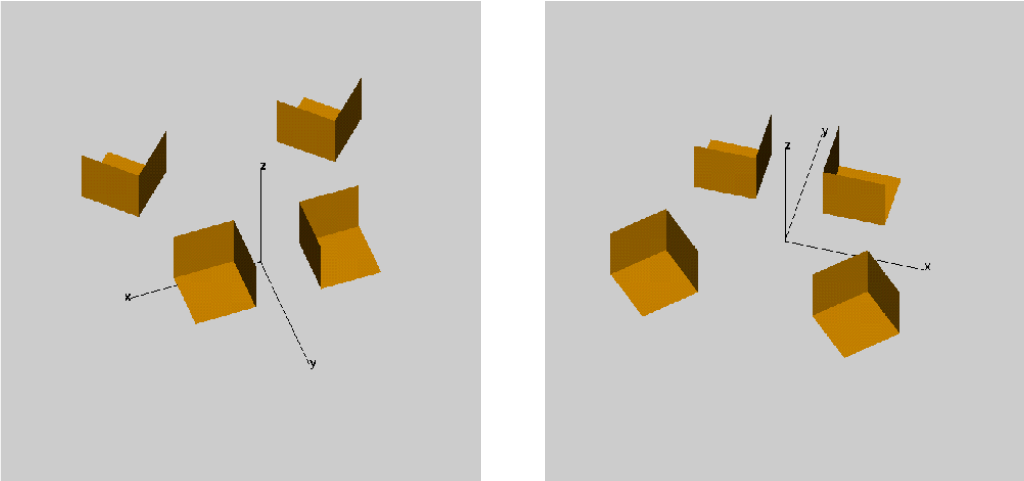


Figure 5.6. Two views of the second multiple-primitive target.

target component	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}\,T}$ [cm]	θ_i^r	$\theta_i^{\mathbf{p}\,T}$ [°]
trihedral 1	1	23.75	$[\ 15.2 \quad 7.6 \quad 15.2\ ]$	—	$[\ 35.26 \quad -45 \quad 0\ ]$
trihedral 2	1	23.75	$[\ -15.2 \quad 7.6 \quad 15.2\ ]$	—	$[\ 35.26 \quad -135 \quad 0\ ]$
trihedral 3	1	23.75	$[\ -45.7 \quad -30.5 \quad 15.2\ ]$	—	$[\ 35.26 \quad 90 \quad 0\ ]$
trihedral 4	1	23.75	$[\ 45.7 \quad -30.5 \quad 15.2\ ]$	—	$[\ 35.26 \quad 90 \quad 0\ ]$

Table 5.14. Parameters of the components of target two.

AVSD	target component	P_{disc}	type confusion	overfit	fraction spurious
10°	trihedral 1	1.000	[1.000 0.000 0.000 0.000]	0.032	0.000
	trihedral 2	1.000	[1.000 0.000 0.000 0.000]		
	trihedral 3	1.000	[1.000 0.000 0.000 0.000]		
	trihedral 4	1.000	[1.000 0.000 0.000 0.000]		
20°	trihedral 1	1.000	[1.000 0.000 0.000 0.000]	0.080	0.000
	trihedral 2	1.000	[1.000 0.000 0.000 0.000]		
	trihedral 3	1.000	[1.000 0.000 0.000 0.000]		
	trihedral 4	1.000	[1.000 0.000 0.000 0.000]		

AVSD	target	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
	component	bias	stdev	$\lVert\text{bias}\rVert$	$\sqrt{\text{tr(cov)}}$	az/el	rot
10°	trihedral 1	−0.720	0.524	5.606	2.321	3.505	17.184
	trihedral 2	−0.826	0.541	5.755	2.393	3.910	18.734
	trihedral 3	−0.685	0.629	7.208	2.686	3.318	21.761
	trihedral 4	−0.747	0.574	7.046	2.891	2.809	18.662
20°	trihedral 1	−1.192	1.192	3.816	4.246	5.812	22.188
	trihedral 2	−1.228	1.008	4.530	4.809	6.187	23.780
	trihedral 3	−0.930	0.981	7.426	4.507	4.519	24.961
	trihedral 4	−0.884	0.778	6.978	4.308	3.758	24.087

Table 5.15. Results from second multiple-primitive-target experiments.

all components of this target are the same type and size, they can be compared to the same single-primitive-target benchmark, *i.e.*, the trihedral benchmark of Table 5.12. The performance of the algorithm on this target is presented in Table 5.15. The statistics for each component are similar; because each component is identical we can combine the error statistics for the four components into a single set of ensemble statistics describing the average error over all four target components. (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 trihedral.) These ensemble statistics are presented in Table 5.16. Comparing these results to the trihedral benchmarks of Table 5.12, we see that most of the error statistics from this experiment are similar to the benchmarks. One effect of the closer proximity of primitives is an increase in the fraction of runs in which the algorithm overestimates model order (3.2% of the 10°-AVSD runs and 8.0% of the 20°-AVSD runs) compared to the experiments of Section 5.2 or Section 5.4.1. Almost all of the overfit runs observed for this target were the result of improper clustering of trihedral double-bounce reports between primitives in the initialization stage, made more likely by the close proximity and symmetric orientation of primitives in this target.

AVSD	P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
			bias	stdev	$\|\text{bias}\|$	$\sqrt{\text{tr(cov)}}$	az/el	rot
10°	1.000	1.000	−0.744	0.568	6.404	2.583	3.408	19.158
20°	1.000	1.000	−1.058	1.001	5.688	4.473	5.163	23.775

Table 5.16. Ensemble statistics for second multiple-primitive-target experiments.

5.5 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 will not be possible. A model generation algorithm will have to contend with unknown dimensions and multiple-sized instances of each primitive class. As described in Section 3.1, there are several ways to deal with this issue, ranging from the method we have chosen (*i.e.*, choosing a specific set of nominal dimensions to construct a single scattering model for all primitives of a single class) 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 to some degree (see Section 2.3.2 and Appendix A). 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 and tophats, whose primary response mechanisms do not vary with primitive dimension. We would thus expect the effects of primitive dimension mismatch to be more severe for dihedrals and cylinders than for trihedrals and tophats. Furthermore, due to the similarity between the dihedral and cylinder shaping functions—both involve a sinc term whose mainlobe width is inversely proportional to primitive height (see Table 2.1)—we would expect the effects of primitive dimension mismatch to be similar for dihedrals and trihedrals. We thus focus on the effects of primitive-dimension mismatch as it concerns dihedrals, since this will be similar to the effects on cylinders and more marked than the effects on tophats and trihedrals.

Table 5.17 lists dimensions of three different 25-dBsm dihedrals, denoted "short," "square," and "tall." To establish a benchmark against which to gauge the effects of size mismatch, we ran 10°-AVSD experiments similar to those of Section 5.2 for each of these three dihedrals, using $S_{\theta_i^t}(\cdot)$ perfectly matched to the true dihedral dimensions in each case. The results of these experiments are presented in Table 5.18. As in Section 5.2, Monte Carlo simulations were continued until the dihedral was discovered and correctly identified 500 times. (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 discovery.) The variation in the

dihedral	dimensions
short	$a = 53.04$ cm, $b = 20.90$ cm
square	$a = 33.30$ cm, $b = 33.30$ cm
tall	$a = 20.90$ cm, $b = 53.04$ cm

Table 5.17. Dimensions of three dihedrals. (The dimension notation used here corresponds to that of Figure 2.10.) These dimensions give a base amplitude of 25 dBsm for each dihedral.

dihedral	P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
			bias	stdev	$\lVert$bias$\rVert$	$\sqrt{\text{tr(cov)}}$	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

Table 5.18. Results from dihedral experiments with perfectly matched nominal and actual dimensions.

statistics between the perfectly matched short, square, and tall dihedrals is attributable to the different responses of these three dihedrals: although they all have the same base amplitude, the specularity of the response increases with dihedral height, so that the short dihedral has a broader out-of-plane response than the tall dihedral, and is thus more observable. This shows up most noticeably in the differences in P_{disc} between the three dihedrals.

Table 5.19 presents results from experiments in which the nominal dimensions used to construct the dihedral $S_{\theta_i^t}(\cdot)$ were mismatched to the dimensions of the primitive in the target scene. In particular, for the experiments summarized in Table 5.19, $S_{\theta_i^t}(\cdot)$ was constructed using nominal dimensions given by the square dihedral ($a = 33.30$ cm, $b = 33.30$ cm). The actual dihedrals used in the experiments, however, were the short or tall dihedrals. Comparing the results of Table 5.19 to those of Table 5.18, it is apparent that this dimension mismatch has a relatively minor effect on performance. (The most noticeable effect is the change in base amplitude bias, induced by the imperfect match between $S_{\theta_i^t}(\cdot)$ and the observed dihedral response.) 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.

dihedral	P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
			bias	stdev	$\lVert$bias$\rVert$	$\sqrt{\text{tr(cov)}}$	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

Table 5.19. Results from dihedral experiments with mismatched nominal and actual dimensions.

Chapter 6

Dealing with Obstruction

THE target model generation algorithm we have developed in the preceding chapters is based on a flexible framework for fusing information extracted from multiple SAR images. This framework is flexible enough to support the use of a different feature extraction method than the one we have chosen here, or a different SAR imaging scenario, or a different set of reflector primitives, or a different target parameterization. The basis of our framework's flexibility is the set of assumptions made in Section 3.3: these assumptions permit construction and use of a wide variety of measurement models, including models that capture and describe physical phenomena that our current model neglects. The capabilities of any algorithm operating in this framework will depend on the physical realism and descriptive utility of its associated measurement model.

In this chapter we address one important shortcoming of our algorithm: its inability to accommodate properly partially obstructed primitives. In particular, our measurement model and algorithm are unable to model the effects of obstruction—namely, the lack of detections of a primitive at obstructed viewing angles. In Section 6.1 we demonstrate the detrimental effects of obstruction on the existing algorithm. In Section 6.2 we consider various approaches to dealing with obstruction, ranging from a sophisticated phenomenological approach that models obstruction directly in terms of its physical basis to a data-driven approach that models obstruction based solely on observed primitive responses. We consider the implications of the different approaches for practical implementation of a new target model generation algorithm, and opt to pursue the data-driven approach, which enables realistic modeling of the effects of obstruction without necessitating significant coupling of primitive computations. In Section 6.3 we describe how such an approach can be implemented with relatively minor modifications to the existing measurement model and algorithm. In Section 6.4 we present results demonstrating the behavior of this modified algorithm on targets containing primitives that are partially obstructed.

■ 6.1 The Effects of Obstruction on the Existing Algorithm

Consider the four single-primitive targets described in Table 6.1. The responses of these four primitives over the 2.5° viewsphere grid (see Section 5.1.1) are depicted in

primitive	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}T}$ [cm]	θ_i^r [cm]	$\theta_i^{\mathbf{P}T}$ [°]
trihedral	1	23.75	[30.5 0 15.2]	—	[35.26 0 0]
tophat	2	10.00	[30.5 0 15.2]	18.39	[90 0 —]
dihedral	3	25.00	[30.5 0 15.2]	—	[25 0 0]
cylinder	4	10.00	[30.5 0 15.2]	17.70	[65 180 —]

Table 6.1. Parameters of four primitives.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tri	1.000	1.000	−0.703	0.541	6.799	2.211	2.890	19.246	—
	top	1.000	1.000	−1.382	0.191	0.409	1.288	1.329	—	1.056
	dih	0.953	0.996	−2.033	2.338	5.764	13.890	11.713	9.281	—
	cyl	0.490	0.979	0.080	1.625	0.410	22.169	1.286	—	6.629
20°	tri	1.000	1.000	−0.865	0.914	6.271	4.744	4.641	25.058	—
	top	1.000	1.000	−1.298	0.365	0.446	2.175	2.388	—	1.730
	dih	0.437	0.961	−3.225	4.423	4.862	23.908	15.660	14.963	—
	cyl	0.109	0.966	0.455	2.771	1.786	30.845	3.433	—	9.980

Table 6.2. Results from unobstructed primitives.

Figure 6.1. These primitives are identical to those used to provide the benchmark results for the experiments of Section 5.4.1; in that section (and throughout Chapter 5) we described how the different observabilities and characteristic responses of each primitive affect the performance of the algorithm. The performance of the algorithm on these four primitives is given in Table 6.2.[1]

Suppose now that each of these four primitives was partially obstructed so that its observed response was not as depicted in Figure 6.1, but instead appeared as in Figure 6.2. How would the performance of the existing algorithm on these primitives change? Clearly, we would expect this obstruction to have some detrimental effect. We could point to two likely sources of performance degradation. First of all, an obstructed primitive will be less observable than an unobstructed primitive—it will produce fewer reports in any data set—so we would expect P_{disc} to decrease and the quality of all estimates to decline. Furthermore, we would expect the gross differences between a primitive's observed response in Figure 6.2 and its assumed canonical response in Figure 6.1 to result in an especially marked increase in the pose and base amplitude errors, even beyond the degradation attributable to the lower observability of each primitive, as the algorithm attempts to orient $S_{\theta_i^t}(\cdot)$ to a set of measurements that is a decidedly imperfect match to that function. These two effects will both have a negative influence on algorithm performance.

To characterize the overall effect of obstruction on the algorithm, we performed

[1]This table is a condensed presentation of the benchmark results presented in Section 5.4.1.

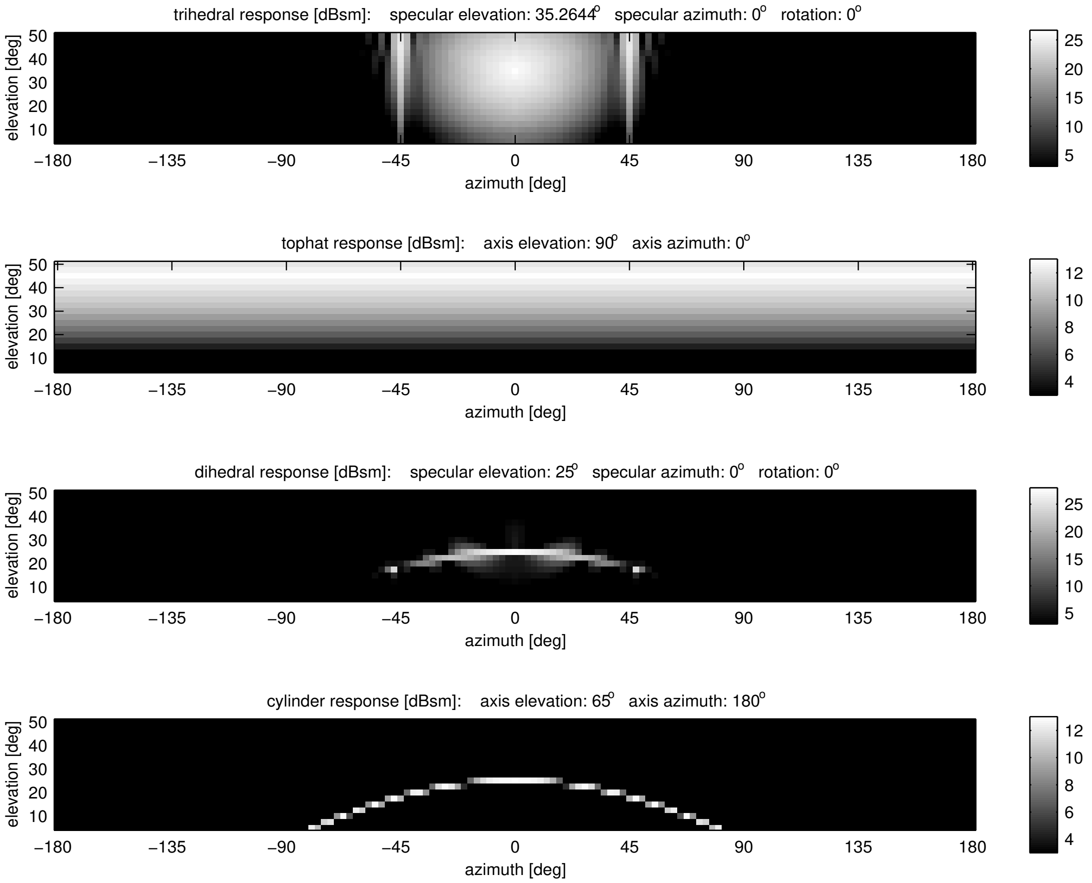


Figure 6.1. Responses of four unobstructed primitives. These primitives are described in Table 6.1. These images depict each primitive's $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ at every point on the 2.5° viewsphere grid. (The images are clipped below 3 dBsm because primitives are not detectable from viewing angles at which $A(\boldsymbol{\theta}_i, \psi_k, \phi_k)$ is too low.)

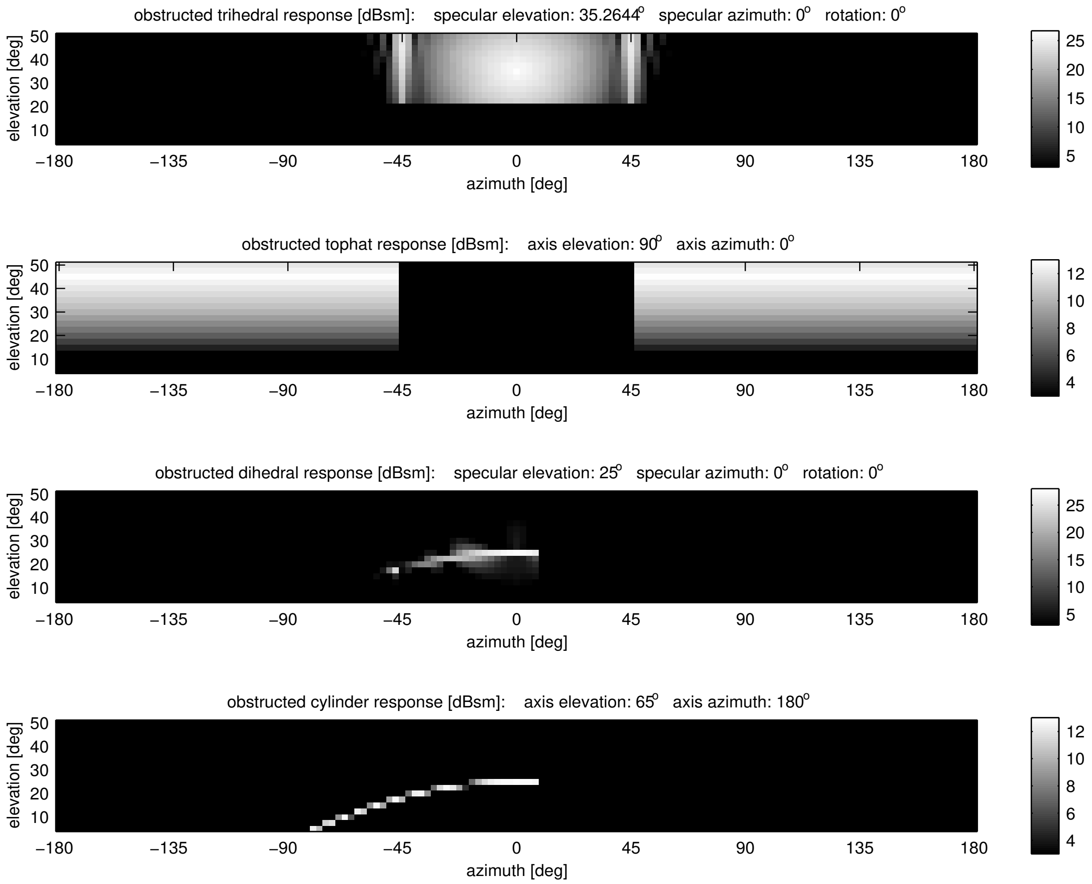


Figure 6.2. Responses of four obstructed primitives. These primitives are the same primitives depicted in Figure 6.1, but with portions of their responses artificially censored to simulate the effects of obstruction. The trihedral has been obstructed at elevations 20° and below, the tophat between azimuth −45° and 45°, and the dihedral and cylinder for azimuths 10° and greater.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse [cm]
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	
10°	tri	1.000	1.000	−1.644	1.294	9.276	3.535	15.949	31.389	—
	top	1.000	1.000	−3.510	0.736	2.295	2.037	12.323	—	1.881
	dih	0.805	1.000	−3.474	3.311	5.030	16.369	18.258	20.621	—
	cyl	0.158	0.889	−0.278	2.037	2.178	19.474	2.412	—	8.079
20°	tri	0.992	0.988	−1.423	1.635	9.790	6.639	16.515	32.324	—
	top	1.000	1.000	−2.718	1.367	3.108	3.884	18.518	—	3.361
	dih	0.183	0.953	−3.926	4.978	7.373	26.608	18.649	23.989	—
	cyl	0.026	0.922	0.499	2.522	2.905	32.413	3.873	—	10.947

Table 6.3. Performance of existing algorithm on partially obstructed primitives. The obstructed responses are depicted in Figure 6.2.

separate experiments on each of the four primitives described in Table 6.1. These experiments were identical to those of Section 5.2[2] except for one important detail: the primitive responses were artificially censored to conform to the obstructed responses of Figure 6.2. In particular, we constructed SAR-image data sets for unobstructed primitives as described in Section 5.1.1, and selected a random subset of images at an AVSD of 10° or 20° for each algorithm run as before. In the formation of **Z** from the image set in any algorithm run, however, we discarded all reports extracted from images whose viewing angles lie in the obstructed regions depicted in Figure 6.2. Thus the feature set **Z** contained no measurements from obstructed viewing angles, but only missed detections. The results from these experiments are presented in Table 6.3.

Comparison of the results of these experiments to the unobstructed-primitive results of Table 6.2 reveals the precise nature of the detrimental effect of unmodeled obstruction on the performance of the algorithm. Almost all performance statistics are worse in the presence of obstruction: P_{disc} values have declined, most error covariances and rmses have increased, many of the pose rmses and base amplitude error covariances have increased dramatically, and the negative base amplitude bias has increased significantly in many cases. Table 6.3 indicates that the effects of obstruction vary for each primitive, and are fundamentally dependent on the observability of the primitive. In particular, obstruction has a different effect in high- and low-observability situations. Limited obstruction is unlikely to affect the discoverability of highly observable primitives (*i.e.*, trihedrals and tophats), which will generally produce enough reports to be discovered even if portions of their responses are not visible; obstruction will, however, have an adverse effect on the pose and base amplitude estimates produced for these primitives, due to the gross mismatch between their observed and assumed responses. Low-observability primitives (*i.e.*, dihedrals and cylinders), on the other hand, generally produce only a handful of reports in any data set even when unobstructed; any

[2] All measurement-model and algorithmic parameters in this and all other experiments of this chapter are as given in Tables 5.2 and 5.3, respectively; the primitive dimensions used to construct the $S_{\theta_i^t}(\cdot)$ for this and all other experiments of this chapter are as given in Table 5.10.

obstruction will tend to have an adverse effect on the discoverability of these primitives, and will be accompanied by a less marked effect on their pose errors.

To clarify the different impact of obstruction on high- and low-observability primitive parameter estimates, we can consider the number of reports that each primitive will generally produce in any data set. In particular, recall that discoverability of any primitive requires that the primitive produce at least n_{min} reports of amplitude a_{min} or greater; furthermore, estimate quality will improve as the number of reports (*i.e.*, primitive observability) increases. Figure 6.3 presents histograms of the number of super-threshold reports (*i.e.*, reports of amplitude a_{min} or greater) produced by each obstructed and unobstructed primitive for a 10°-AVSD data set.[3] Figure 6.4 presents similar histograms for a 20°-AVSD data set. Each of these figures contains four plots, one for each primitive type. Each plot in turn contains two histograms: the dashed line corresponds to the number of super-threshold reports produced by the unobstructed primitive, and the solid line corresponds to the number of super-threshold reports produced by the obstructed primitive. The vertical dotted line in each plot corresponds to n_{min} for that data set. (Recall from Table 5.3 that $n_{\text{min}} = 4$ for 10°-AVSD data sets and $n_{\text{min}} = 3$ for 20°-AVSD data sets.) Obstruction compresses each primitive's histogram toward the left of the plot. We can summarize the effect of obstruction on each histogram as it relates to discoverability and observability by tabulating two quantities from each histogram: $\tilde{P}_{\text{disc}}^{\text{max}}$, the fraction of the histogram to the right of the dotted line (*i.e.*, the fraction of runs in which at least n_{min} super-threshold reports will be available), and $\bar{n}_{\text{disc}}$, the expected number of such reports conditional on discovery. The values of $\tilde{P}_{\text{disc}}^{\text{max}}$ and $\bar{n}_{\text{disc}}$ for each obstructed and unobstructed primitive are listed in Table 6.4. The obstructed trihedral and tophat histograms lie almost entirely to the right of the dotted line, so $\tilde{P}_{\text{disc}}^{\text{max}}$ remains near unity; the quality of the estimates produced for these primitives suffers, however, because $\bar{n}_{\text{disc}}$ decreases. Conversely, obstruction causes a much greater fraction of the dihedral and cylinder histograms to lie to the left of the dotted line, so $\tilde{P}_{\text{disc}}^{\text{max}}$ decreases markedly for these primitives; however, $\bar{n}_{\text{disc}}$ is close to n_{min} for both primitives regardless of obstruction, thus there is much less effect on estimate quality.

As indicated previously, obstruction affects algorithm performance through two distinct channels: first, by reducing primitive observability as just described, and second, by introducing a mismatch between the observed primitive responses and the assumed primitive responses. In theory it might be possible to design an algorithm to correct for the second source of error, but there is fundamentally no way to counteract the first source of error without being provided additional measurements. To separate the contributions of these two error sources, we performed a second set of experiments on the obstructed primitives of Figure 6.2 in which the algorithm was provided perfect knowledge of the viewing angles at which obstruction occurs. The optimal way to deal

[3]These histograms were generated in a manner similar to that used to obtain bounds on discoverability in Section 5.2, *i.e.*, by using the primitive responses of Figure 6.1 together with the $P_D(\cdot)$ function of Figure 5.2 to simulate the detection of a primitive in images at randomly selected viewing angles.

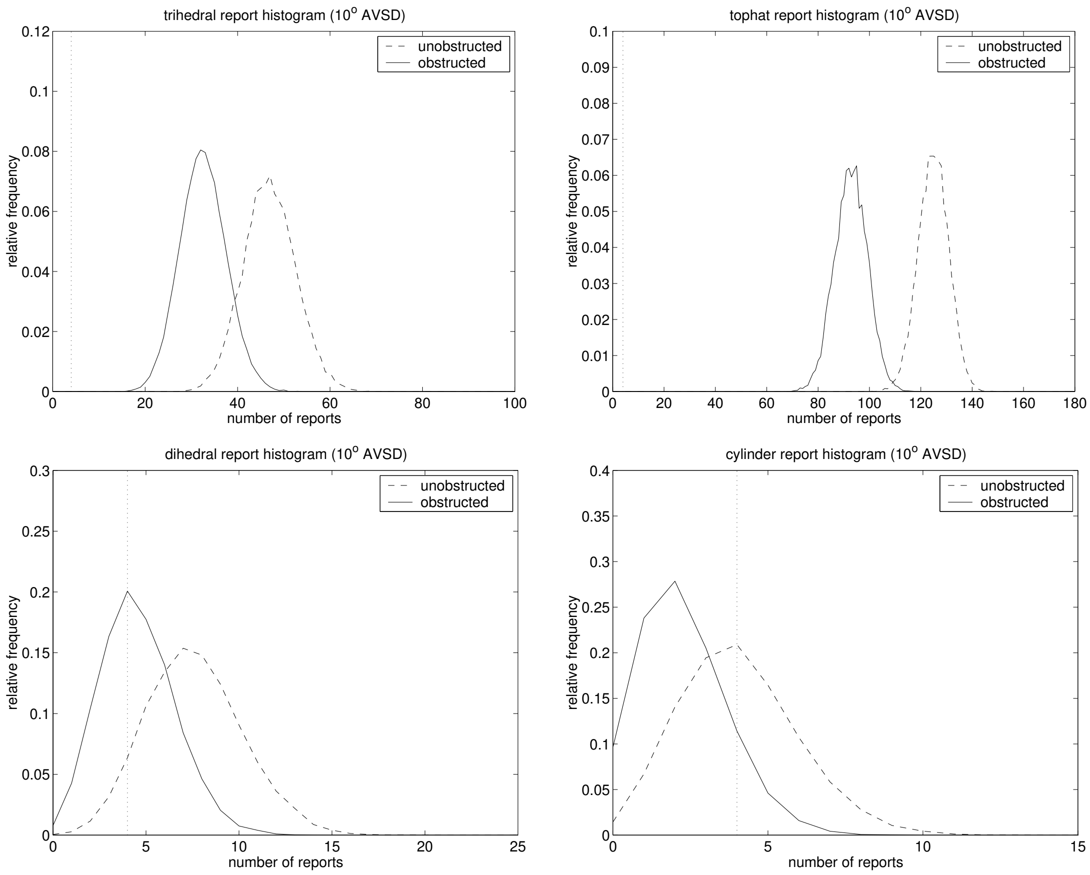


Figure 6.3. Number of reports produced in 10°-AVSD data sets. These four plots (upper left: trihedral; upper right: tophat; lower left: dihedral; lower right: cylinder) are histograms of the number of reports of amplitude $a_{\min}$ or greater produced for each primitive. The dashed line in each plot is the histogram for the unobstructed primitive (see Figure 6.1), the solid line is the histogram for the obstructed primitive (see Figure 6.2), and the vertical dotted line corresponds to $n_{\min}$, the minimum number of reports required for an initialization. Note that the horizontal and vertical scales differ from plot to plot.

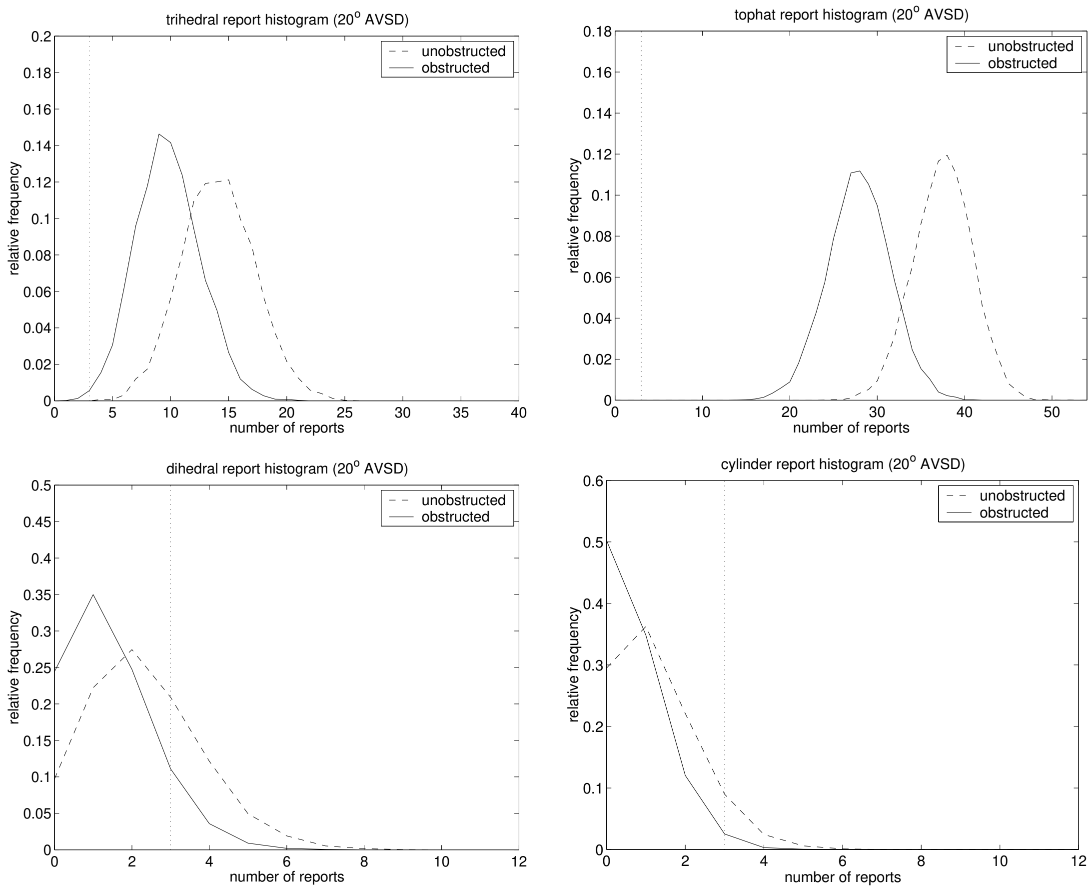


Figure 6.4. Number of reports produced in 20°-AVSD data sets. These four plots (upper left: trihedral; upper right: tophat; lower left: dihedral; lower right: cylinder) are histograms of the number of reports of amplitude $a_{\min}$ or greater produced for each primitive. The dashed line in each plot is the histogram for the unobstructed primitive (see Figure 6.1), the solid line is the histogram for the obstructed primitive (see Figure 6.2), and the vertical dotted line corresponds to $n_{\min}$, the minimum number of reports required for an initialization. Note that the horizontal and vertical scales differ from plot to plot.

primitive		10° AVSD		20° AVSD	
		$\tilde{P}_{\text{disc}}^{\max}$	$\bar{n}_{\text{disc}}$	$\tilde{P}_{\text{disc}}^{\max}$	$\bar{n}_{\text{disc}}$
trihedral	unobstructed	1.000	46.904	1.000	14.099
	obstructed	1.000	32.541	0.999	9.795
tophat	unobstructed	1.000	124.886	1.000	37.501
	obstructed	1.000	92.796	1.000	27.910
dihedral	unobstructed	0.954	7.839	0.406	3.760
	obstructed	0.682	5.586	0.158	3.394
cylinder	unobstructed	0.583	5.297	0.121	3.329
	obstructed	0.182	4.529	0.029	3.142

Table 6.4. Effect of obstruction on number of reports produced by each primitive. The values in this table are compiled from Figures 6.3 and 6.4. Here $\tilde{P}_{\text{disc}}^{\max}$ represents the fraction of runs in which at least $n_{\min}$ reports of amplitude $a_{\min}$ or greater are produced; $\bar{n}_{\text{disc}}$ is the average number of such reports conditional on there being at least $n_{\min}$ available.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tri	1.000	1.000	−0.670	0.516	9.261	3.444	2.808	16.500	—
	top	1.000	1.000	−1.216	0.265	0.155	1.468	1.914	—	1.141
	dih	0.793	0.997	−2.356	3.177	5.476	17.287	15.116	15.188	—
	cyl	0.169	0.955	0.758	2.456	2.530	19.358	1.621	—	8.033
20°	tri	0.992	0.988	−0.837	0.992	9.745	6.588	6.085	27.121	—
	top	1.000	1.000	−1.064	0.461	0.142	2.723	3.460	—	2.156
	dih	0.189	0.916	−3.673	4.971	8.095	25.185	17.970	20.907	—
	cyl	0.026	0.942	0.871	3.306	2.438	36.236	3.449	—	11.906

Table 6.5. Performance of algorithm with perfect knowledge on partially obstructed primitives. The obstructed responses are depicted in Figure 6.2. These results represent the performance of an algorithm given complete knowledge of the viewing angles at which obstruction occurs.

with obstruction in such a case is simply to ignore images in which obstruction is known to have occurred, because these images can provide no information about the primitive. All other images should then be used to form an estimate exactly as before. The results of these "perfect-knowledge" experiments are given in Table 6.5. The degradation in performance observed between Tables 6.2 and 6.5 thus represents the effects of the reduction in primitive observability, while the degradation in performance observed between Tables 6.5 and 6.3 represents the effects of the unmodeled mismatch between observed (obstructed) and assumed primitive responses.

Comparison of Tables 6.2, 6.3, and 6.5 indicates that essentially all of the reduction in discoverability for obstructed primitives is due to lower observability. On the other hand, most of the increase in the tophat and trihedral pose estimate errors is attributable to the response mismatch. The increase in the pose errors of the dihedral and cylinder are attributable to both factors. The increased negative bias observed in the base amplitude estimates of most primitives is attributable to the response mismatch. The slight increase in location and radius-of-curvature estimate errors is chiefly attributable to the lower observability. Comparison of these tables thus indicates that while it would not be possible to design an algorithm to improve detectability or location error statistics in the presence of obstruction, we could improve the quality of pose and base amplitude estimates. These tables also indicate that an algorithm able to deal with obstruction would offer a greater potential benefit to tophats and trihedrals than to dihedrals and cylinders, since the parameter estimates of the latter two low-observability primitives suffer much less from response mismatch than do the parameter estimates of the high-observability trihedral and tophat.

We stress that Table 6.5 represents a benchmark for the best performance we could hope to attain from *any* algorithm faced with the obstructed responses of Figure 6.2, as this performance corresponds to perfect knowledge of the obstructed regions of the viewsphere for each primitive. Our goal in this chapter is thus to design an algorithm that can infer obstruction and correct the error introduced by the unmodeled mismatch between observed obstructed responses and the assumed primitive responses. Such an algorithm, if successful, would approach the performance of Table 6.5 even without any prior knowledge of the presence or extent of obstruction for any target.

6.2 Possible Conceptual Approaches

Conceptually, any algorithm that is able to accommodate the effects of obstruction must do two things: it must infer at which viewing angles a primitive is obstructed, and it must incorporate this knowledge into its estimates of the primitive's descriptive parameters. As indicated previously, the optimal way of incorporating knowledge of obstruction into the estimate of $\boldsymbol{\theta}_i$ is simply to disregard views at which primitive i is known to be obstructed. Because the absence of a report attributable to primitive i in any image influences the estimate of $\boldsymbol{\theta}_i$ only through a missed-detection penalty (see (4.9)), disregarding a known obstructed view is equivalent simply to not penalizing for a missed detection at that view. The conceptual challenge in accommodating obstruction, then, is not the problem of how to correct its effects if it is known, but the more fundamental problem of how to infer its presence and extent. In this section we consider how this might be accomplished.

The approach to dealing with obstruction that is most physically realistic is to couple all primitives so that the detectability of any primitive i in any image depends not only on $\boldsymbol{\theta}_i$ but on the entire vector $\boldsymbol{\theta}$. The existing probability-of-detection function $P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))$ would be replaced with a modified function $\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k)$ that takes into account the relative locations of all primitives in its determination of the detectability of primitive i. For instance, if the location components of $\boldsymbol{\theta}$ indicated that primitive i was downrange from another primitive along the line-of-sight vector defined by ψ_k and ϕ_k, the modified probability-of-detection function would satisfy

$$\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k) \approx 0. \tag{6.1}$$

If $\boldsymbol{\theta}$ indicated that the line of sight between the sensor and primitive i was clear, then we would instead have

$$\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k) \approx P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k)). \tag{6.2}$$

In theory, $\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k)$ could take into account not only the locations, but also the poses and dimensional information (such as provided by θ_i^r and θ_i^a) of each primitive in its determination of the detectability of any individual primitive from any viewing angle. Regardless of whether nonlocational information was used to infer obstruction, however, this approach represents a significant relaxation of Assumption 2 (see Section 3.3), which stated that the detectability of the ith primitive depends only on $\boldsymbol{\theta}_i$ and on the viewing angle.

This approach, while enabling a description of obstruction directly in terms of the processes that are responsible for it, would have important ramifications for the practical implementation of the algorithm. In particular, relaxation of Assumption 2 to allow a coupled probability-of-detection function would introduce significant complications. Although the E step could proceed largely as before (*i.e.*, as a set of probability computations decoupled across images and facilitated by the gating assumption), the M step would be much more difficult: it would no longer be separable into N unrelated maximizations. Instead, maximization over any $\boldsymbol{\theta}_i$ would require joint consideration of all components of $\boldsymbol{\theta}$. Modification of any of these components could affect the detectability of primitive i, and modification of $\boldsymbol{\theta}_i$ could potentially affect the detectability of all other primitives. Implementation of the M step would thus be significantly more complicated than in the present algorithm.

If we wished to maintain the physical sophistication of the above approach (*i.e.*, the dependence of primitive detectability on the locations of all other primitives) and yet continue to perform the M step as a series of N decoupled maximizations, we could make the following approximation: we could simply neglect the dependence of any primitive's detectability on the variation in other primitives' parameters within a single M step. In other words, when obtaining an M-step estimate $\boldsymbol{\theta}_i^{[n+1]}$ we could make the assumption that

$$\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k) \approx \tilde{P}_D(i, \tilde{\boldsymbol{\theta}}(i, n), \psi_k, \phi_k), \tag{6.3}$$

where

$$\tilde{\boldsymbol{\theta}}(i,n) = \begin{bmatrix} \boldsymbol{\theta}_1^{[n]} \\ \vdots \\ \boldsymbol{\theta}_{i-1}^{[n]} \\ \boldsymbol{\theta}_i \\ \boldsymbol{\theta}_{i+1}^{[n]} \\ \vdots \\ \boldsymbol{\theta}_N^{[n]} \end{bmatrix}. \tag{6.4}$$

This would enable $\boldsymbol{\theta}_i^{[n+1]}$ to be obtained without simultaneous maximization over all other primitives' parameter vectors. (The E step could still use the completely coupled $\tilde{P}_D(i, \boldsymbol{\theta}, \psi_k, \phi_k)$ in its determination of the report-to-primitive correspondence probabilities.) As long as the changes in $\boldsymbol{\theta}_i^{[n]}$ were small between iterations, (6.3) might be a reasonable assumption and would enable description of obstruction in terms of its physical basis without necessitating a significant increase in the computational burden of the M step.

The above approach—whether implemented with the simplifying assumption of (6.3) or not—is based on a physically derived model of obstruction. It describes obstruction not only in terms of the obstructed primitive, but also the obstructing primitive. While this approach has the conceptual benefit of ensuring a self-consistent description of obstruction in the context of any target model, it is unnecessarily complicated. In particular, in the context of forming an estimate of any obstructed target primitive, any description of the *obstructing* primitive is completely superfluous. All that is required to correct the effects of obstruction is to identify the views at which obstruction occurs. Thinking of obstruction in terms of both an obstructed and an obstructing primitive is thus beneficial only if it aids in this inferential process.

A different possibility for dealing with obstruction is to infer its presence not from the locations of other primitives, but from the measurements extracted from the SAR images. In other words, we could try to infer the presence and extent of obstruction from $\mathbf{Z}$ instead of $\boldsymbol{\theta}$. Such a data-driven approach would be less physically realistic than the phenomenologically driven approach described above, but would avoid coupling primitive estimation in the M step. In particular, if obstruction for each primitive was inferred separately from $\mathbf{Z}$ without being influenced by the other components of $\boldsymbol{\theta}$, then the M step of the algorithm could proceed as N decoupled maximizations without the aid of any simplifying assumptions. Primitive detectability could still be modeled as a function only of $\boldsymbol{\theta}_i$ and viewing angle, possibly in addition to $\mathbf{Z}$ (or to some related quantities determined by $\mathbf{Z}$, $\boldsymbol{\theta}_i$, and the viewing angle).

Comparison of the obstructed responses of Figure 6.2 to the assumed canonical responses of Figure 6.1 suggests that a data-driven approach might be successful in dealing with obstruction: the obstructed responses of Figure 6.2 retain much of their essential canonical character and are certainly distinguishable and orientable to the human eye. Depending on how we modify our measurement model and algorithm, these responses might also be made orientable to the target model generation algorithm. In the following section we describe our methodology for expanding our existing algorithm to include a data-driven approach to correcting the effects of obstruction.

6.3 Methodology

As indicated previously, in order to accommodate obstruction we must infer at which viewing angles it occurs, and then avoid penalizing missed detections at these viewing angles. This will require some modification to the measurement model of Section 3.3, which will result in a modification $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ and, in turn, necessitate a modification to our existing E and M steps (see Sections 4.2 and 4.3). The precise nature of these modifications will depend on how we incorporate obstruction into the existing framework developed in previous chapters. In this section we propose a specific modification to the measurement model that is well-matched to the structure of the existing algorithm. In particular, this modification describes obstruction in terms of a set of hidden variables that can be naturally accommodated into the EM framework, in a manner similar to that of the vector $\boldsymbol{\lambda}$. These hidden obstruction variables, like the label parameters, not only enable a parsimonious description of the phenomenon they represent, but also enable a description of that obstruction to be efficiently estimated in the EM framework.

Our hidden obstruction variables take the form of NK binary indicators $\beta_{i,k}$, each indicating whether a particular primitive is obstructed from the perspective of a particular image. In particular, we define

$$\beta_{i,k} = \begin{cases} 0, & \text{primitive } i \text{ is not obstructed at viewing angle } (\psi_k, \phi_k), \\ 1, & \text{primitive } i \text{ is obstructed at viewing angle } (\psi_k, \phi_k). \end{cases} \tag{6.5}$$

These variables are clearly not observable, but would offer a concise description of the extent of obstruction for any primitive. Furthermore, if these variables could be observed, they would enable direct correction of the effects of obstruction: we would simply avoid assigning a missed-detection penalty to primitive i in image k whenever $\beta_{i,k} = 1$. As with the label parameters $\lambda_{k,j}$, we define notation that will allow us to refer to the set of all hidden obstruction variables in any image k:

$$\boldsymbol{\beta}_k = \begin{bmatrix} \beta_{1,k} \\ \vdots \\ \beta_{N,k} \end{bmatrix}. \tag{6.6}$$

Similarly, let us define a vector $\boldsymbol{\beta}$ incorporating all obstruction parameters from all images:

$$\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_K \end{bmatrix}. \tag{6.7}$$

Observation of $\boldsymbol{\beta}$ would thus provide complete knowledge of the extent of obstruction for all primitives in all images.

In order to use the hidden parameters $\beta_{i,k}$ to describe and correct the effects of obstruction, we must incorporate $\boldsymbol{\beta}$ into our measurement model. We propose a new measurement model in the following section. We then show how this new model affects the expected-log-likelihood function $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ and describe its implications are for the E and M steps of the algorithm, as well as the initialization.

6.3.1 Modification to the Measurement Model

We require specification of a new measurement model $p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta})$ that includes a complete probabilistic description of the relationship between the newly introduced obstruction parameters $\boldsymbol{\beta}$ and the previously defined quantities $\boldsymbol{\lambda}, \mathbf{Z}$, and $\boldsymbol{\theta}$. We can decompose $p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta})$ as

$$p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta}) = p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\theta})\, p(\boldsymbol{\lambda}, \boldsymbol{\beta}|\boldsymbol{\theta}). \tag{6.8}$$

The $p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\theta})$ component of (6.8) is analogous to the conditional fine-level uncertainty model of Section 3.3, and $p(\boldsymbol{\lambda}, \boldsymbol{\beta}|\boldsymbol{\theta})$ is analogous to the coarse-level uncertainty model of that section, and is related to the identity of each measurement. We will assume the same fine-level uncertainty model as was used in Chapter 3, *i.e.*, in which the component measurements of any report $\mathbf{Z}_{k,j}$ depend only on the identity of that report. Because $\boldsymbol{\beta}$ offers no further information than $\boldsymbol{\lambda}$ about the identity of each report, we may thus write

$$\begin{aligned} p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\beta}, \boldsymbol{\theta}) &= p(\mathbf{Z}|\boldsymbol{\lambda}, \boldsymbol{\theta}) \\ &= \prod_{k=1}^{K} p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta}), \end{aligned} \tag{6.9}$$

where the $p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta})$ terms are exactly as specified in (3.40).

The coarse-level uncertainty component $p(\boldsymbol{\lambda}, \boldsymbol{\beta}|\boldsymbol{\theta})$ in (6.8) can be decomposed as

$$p(\boldsymbol{\lambda}, \boldsymbol{\beta}|\boldsymbol{\theta}) = p(\boldsymbol{\lambda}|\boldsymbol{\beta}, \boldsymbol{\theta})\, p(\boldsymbol{\beta}|\boldsymbol{\theta}). \tag{6.10}$$

Unlike the fine-level uncertainty of (6.9), the coarse-level uncertainty does change with the introduction of $\boldsymbol{\beta}$. In particular, the inclusion of $\boldsymbol{\beta}$ imposes a new constraint on $\boldsymbol{\lambda}$: by definition, if primitive i is obstructed in image k, then it cannot produce a report in that image. Thus $\lambda_{k,j}$ can take on value i only if $\beta_{i,k} = 0$. There is thus a dependence of $\boldsymbol{\Lambda}$ (the set of all possible $\boldsymbol{\lambda}$) on any particular choice of $\boldsymbol{\beta}$ in the conditional pmf $p(\boldsymbol{\lambda}|\boldsymbol{\beta},\boldsymbol{\theta})$.

To enable specification of $p(\boldsymbol{\lambda}|\boldsymbol{\beta},\boldsymbol{\theta})$, we make several basic assumptions that are analogues to those of Section 3.3, but which have been slightly modified to accommodate $\boldsymbol{\beta}$. In particular, we assume that the detectability of the ith primitive in image k depends only on $\boldsymbol{\theta}_i$, $\beta_{i,k}$, and on the image viewing angle; we assume that missed detections are conditionally independent from image to image given $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$; finally, we assume that false alarms are independent from image to image and are independent of missed detections. These assumptions allow us to write

$$p(\boldsymbol{\lambda}|\boldsymbol{\beta},\boldsymbol{\theta}) = \prod_{k=1}^{K} p(\boldsymbol{\lambda}_k|\boldsymbol{\beta}_k,\boldsymbol{\theta}), \tag{6.11}$$

so that the identities $\boldsymbol{\lambda}_k$ of all reports in any image k can be expressed in terms of $\boldsymbol{\theta}$ and the obstruction parameters $\boldsymbol{\beta}_k$ for that image, without any dependence on terms from other images. (This is analogous to the decomposition of $p(\boldsymbol{\lambda}|\boldsymbol{\theta})$ into independent terms for each image in (3.20).)

To derive an expression for $p(\boldsymbol{\lambda}_k|\boldsymbol{\beta}_k,\boldsymbol{\theta})$ we must specify false-alarm and primitive-detectability models that include $\boldsymbol{\beta}_k$. If $\beta_{i,k} = 1$ then, by definition, primitive i is not detectable in image k and there should thus be no penalty for a missed detection. If $\beta_{i,k} = 0$, we will model the detectability of primitive i in image k as $P_{D}{}'_{k,i}$, exactly as in Section 3.3. We will continue to model false alarms as a Poisson arrival process, and to assume that there is no preferential or systematic ordering of the elements of $\boldsymbol{\lambda}_k$. We may thus write

$$p(\boldsymbol{\lambda}_k|\boldsymbol{\beta}_k,\boldsymbol{\theta}) = \frac{e^{-\gamma_{FA}V}(\gamma_{FA}V)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} \left(1 - P_{D}{}'_{k,i}\right)^{1-\beta_{i,k}} \cdot \prod_{j:\lambda_{k,j}\neq 0} \frac{P_{D}{}'_{k,\lambda_{k,j}}}{1 - P_{D}{}'_{k,\lambda_{k,j}}}. \tag{6.12}$$

Note that the only functional difference between $p(\boldsymbol{\lambda}_k|\boldsymbol{\beta}_k,\boldsymbol{\theta})$ in (6.12) and $p(\boldsymbol{\lambda}_k|\boldsymbol{\theta})$ as expressed in (3.32) is the $1 - \beta_{i,k}$ exponent on the missed-detection penalty in (6.12).[4] This accomplishes the proper incorporation of obstruction knowledge, *i.e.*, the avoidance of applying a missed-detection penalty at viewing angles at which a primitive is known to be obstructed.

Complete specification of the new coarse-level uncertainty model $p(\boldsymbol{\lambda},\boldsymbol{\beta}|\boldsymbol{\theta})$ now lacks only an expression for $p(\boldsymbol{\beta}|\boldsymbol{\theta})$. As indicated in Section 6.2, we wish to maintain the current decoupled nature of the M step, which can be separated into independent maximizations for each primitive. This imposes the constraint that $p(\boldsymbol{\beta}|\boldsymbol{\theta})$ should be specified as

$$p(\boldsymbol{\beta}|\boldsymbol{\theta}) = \prod_{i=1}^{N} p(\beta_{i,1},\ldots,\beta_{i,K}|\boldsymbol{\theta}_i). \tag{6.13}$$

[4]As described previously, the conditioning on $\boldsymbol{\beta}_k$ also imposes implicit restrictions on $\boldsymbol{\Lambda}_k$, although this is not expressed explicitly in (6.12).

This product represents a prior distribution for patterns of obstruction on the viewsphere for any primitive.

Intuitively, we would expect obstruction to occur at proximate angles on the viewsphere. For instance, given a fixed number of obstructed views in some data set, it seems more plausible that these views would be from similar angles than from angles spread widely across the viewsphere. In other words, an intuitively plausible prior would reward localization of obstructed views. This suggests the use of a Markov random field (MRF) model [37,116] for $p(\boldsymbol{\beta}|\boldsymbol{\theta})$, in which the prior probability of obstruction at any viewing angle is specified conditionally on the probabilities of obstruction in the neighborhood of that viewing angle. Such an approach, while intuitively and physically reasonable, creates several difficulties. First is the adaptation of standard MRF models, which are usually applied to planar lattices, to sparse and unequally spaced samples on the sphere. Second is the effect of an MRF-type model on the implementation of the E step. In the existing algorithm, the computation of the $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ is made tractable by the full decoupling of the measurement model across views, *i.e.*, by the ability to express the measurement model as a product of K independent terms. Unless we choose a prior $p(\boldsymbol{\beta}|\boldsymbol{\theta})$ that also decouples across images, there will be a combinatorial explosion in the computational burden associated with the E step.

We model $p(\boldsymbol{\beta}|\boldsymbol{\theta})$ using a simple prior in which the probability of obstruction of a primitive in any image is constant and independent of the obstruction of the primitive in all other images. Specifically, we model

$$p(\boldsymbol{\beta}|\boldsymbol{\theta}) = \prod_{k=1}^{K} \prod_{i=1}^{N} P_b^{\beta_{i,k}} (1 - P_b)^{1-\beta_{i,k}}, \tag{6.14}$$

where P_b is a parameter representing the probability of obstruction of any primitive in any image. (We discuss appropriate selection of P_b in Section 6.3.5.) This model, while less physically meaningful than other possibilities as described above, has the important benefit of allowing the E step to proceed in its present form, *i.e.*, as a computation decoupled across images.

We can use the previous equations to present a full specification of the modified measurement model:

$$\begin{aligned} p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta}) = \prod_{k=1}^{K} \Bigg[& p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta}) \cdot \frac{e^{-\gamma_{FA} V} (\gamma_{FA} V)^{F_k}}{M_k!} \cdot \prod_{j:\lambda_{k,j} \neq 0} \frac{P_{D'_{k,\lambda_{k,j}}}}{1 - P_{D'_{k,\lambda_{k,j}}}} \\ & \cdot \prod_{i=1}^{N} P_b^{\beta_{i,k}} \left((1 - P_{D'_{k,i}})(1 - P_b) \right)^{1-\beta_{i,k}} \Bigg], \end{aligned} \tag{6.15}$$

where $p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta})$ is specified in (3.40). Note the similarity of (6.15) to the original measurement model of Section 3.3: the only differences are the presences of the P_b and $1 - P_b$ terms, the binary $\beta_{i,k}$ and $1 - \beta_{i,k}$ exponents, and the implicit restriction of $\boldsymbol{\Lambda}_k$ to accommodate additional constraints imposed by $\boldsymbol{\beta}_k$.

■ 6.3.2 Modification to the E Step

Modification of the measurement model leads to a change in the expected-log-likelihood function that must be computed in each E step. In particular, the inclusion of the hidden variables $\boldsymbol{\beta}$ into the measurement model indicates that the expected-log-likelihood is now of the form

$$
\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) &= E\left[\log p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta}) \,\middle|\, \mathbf{Z}, \boldsymbol{\theta}^{[n-1]}\right] \\
&= \sum_{(\boldsymbol{\lambda},\boldsymbol{\beta})\in(\boldsymbol{\Lambda},\mathbf{B})} \left[\log p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta})\right] p(\boldsymbol{\lambda}, \boldsymbol{\beta}|\mathbf{Z}, \boldsymbol{\theta}^{[n-1]}),
\end{aligned}
\tag{6.16}
$$

where $\mathbf{B}$ is the set of all possible $\boldsymbol{\beta}$, *i.e.*, a set of 2^{KN} binary vectors each corresponding to a distinct choice of $\boldsymbol{\beta}$, and $(\boldsymbol{\Lambda}, \mathbf{B})$ is the joint set of all possible $(\boldsymbol{\lambda}, \boldsymbol{\beta})$ pairs. A derivation similar to that of Appendix B shows that with N specified, the expected-log-likelihood can still be written as

$$
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_{i=1}^{N} Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) + C_K, \tag{6.17}
$$

where each of the N terms $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ depends only on primitive i and can be further decomposed into a sum of K terms, one for each image:

$$
Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{k=1}^{K} Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}), \tag{6.18}
$$

where

$$
\begin{aligned}
Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = \sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) &\left[\log \frac{P_{D'_{k,i}}}{1 - P_{D'_{k,i}}} - \frac{1}{2\sigma_a^2} (a_{k,j} - A(\boldsymbol{\theta}_i, \psi_k, \phi_k))^2 \right. \\
&\left. + \log \rho'_{k,j} - \frac{1}{2} (\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i))^T \mathbf{R}^{-1} (\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)) \right] \\
&+ (1 - \Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})) \log(1 - P_{D'_{k,i}}),
\end{aligned}
\tag{6.19}
$$

or, equivalently,

$$
\begin{aligned}
Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]}) = {} & \Pr(\delta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) \log {P_D}'_{k,i} \\
& + \Big[1 - \Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) - \Pr(\delta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\Big] \log\left(1 - {P_D}'_{k,i}\right) \\
& + \sum_{j=1}^{M_k} \Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) \Bigg[-\frac{1}{2\sigma_a^2}\left(a_{k,j} - A(\boldsymbol{\theta}_i,\psi_k,\phi_k)\right)^2 \\
& \qquad + \log \rho'_{k,j} - \frac{1}{2}\left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)^T \mathbf{R}^{-1} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)\Bigg].
\end{aligned}
\tag{6.20}
$$

This bears a striking resemblance to the expected-log-likelihood function of Section 4.2: the only difference is the presence of the subtractive $\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ terms above. These terms have exactly the desired effect: they avoid penalizing for a missed detection in an image from which the probability of obstruction has been determined to be significant.

As before, the E step requires calculation of $\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ terms for all (i,j,k) triplets; in addition, complete determination of the modified $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ also requires calculation of $\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ for all (i,k) pairs. These computations can proceed by applying Bayes' rule to $p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k)$. In particular, we can write

$$
\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) = \frac{\sum_{(\boldsymbol{\Lambda}_k,\mathbf{B}_k):\lambda_{k,j}=i} p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}{\sum_{(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k)\in(\boldsymbol{\Lambda}_k,\mathbf{B}_k)} p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}, \tag{6.21}
$$

where $\mathbf{B}_k$ is the set of all possible $\boldsymbol{\beta}_k$ and $(\boldsymbol{\Lambda}_k,\mathbf{B}_k)$ is the set of all possible $(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k)$ pairs. We can obtain $\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ by similar means:

$$
\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) = \frac{\sum_{(\boldsymbol{\Lambda}_k,\mathbf{B}_k):\boldsymbol{\beta}_{i,k}=1} p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}{\sum_{(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k)\in(\boldsymbol{\Lambda}_k,\mathbf{B}_k)} p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k|\boldsymbol{\theta}^{[n]})}. \tag{6.22}
$$

As in Section 4.2, these computations can be greatly simplified by gating.

Note that due to the required consideration of the 2^N-element set $\mathbf{B}_k$ in the summations of (6.21) and (6.22), it would appear that inclusion of $\boldsymbol{\beta}$ in the measurement model has increased the computational burden of the E step by a factor of 2^N. In fact, the actual increase in computational complexity is much smaller, because evaluation of $p(\boldsymbol{\lambda}_k,\boldsymbol{\beta}_k,\mathbf{Z}_k|\boldsymbol{\theta}^{[n]})$ for different values of $\boldsymbol{\beta}_k$ requires many of the same calculations. In practice the $\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ and $\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ terms can be obtained without much more difficulty than was required to obtain the $\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})$ in Section 4.2. This is an additional benefit of the simple prior (6.14) we have assumed for $\boldsymbol{\beta}$.

6.3.3 Modification to the M Step

As before, the M step entails maximization of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ over all $\boldsymbol{\theta}$. Because of the similarity of the modified $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ to the original, the M step can proceed largely as before. The single difference is the inclusion of the subtractive $\Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ term in the missed-detection penalty term of (6.20). Because this change affects only the pose- and base-amplitude-dependent terms of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$, the ECM estimates of $\theta_i^{\mathbf{x}}$ and θ_i^r can be obtained by maximization of (4.25) exactly as before. The estimates of $\theta_i^{\mathbf{P}}$ and θ_i^a must be obtained by maximizing a slightly modified version of (4.24):

$$\sum_{k=1}^{K}\left(\sum_{j=1}^{M_k}\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\left[\log\frac{P_D{}'_{k,i}}{1-P_D{}'_{k,i}} - \frac{1}{2\sigma_a^2}\left(a_{k,j}-A(\boldsymbol{\theta}_i,\psi_k,\phi_k)\right)^2\right]\right.$$
$$\left. + \left[1-\Pr(\beta_{i,k}=1|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\right]\log(1-P_D{}'_{k,i})\right). \tag{6.23}$$

Implementation of the M step is thus not substantially changed by the inclusion of $\boldsymbol{\beta}$ in the measurement model.

6.3.4 Modification to the Initialization Procedure

The initialization procedure described in Section 4.4 is based on an agglomerative clustering algorithm that provides an estimate of $\boldsymbol{\lambda}$ from which to form $\boldsymbol{\theta}^{[0]}$. The presence of the $\Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ terms in (6.23) indicates that the initialization procedure must be modified to provide some estimate of $\boldsymbol{\beta}$ prior to the construction of $\boldsymbol{\theta}^{[0]}$. A simple approach would be to neglect the effects of obstruction in the initialization. In other words, we could assume that all $\beta_{i,k} = 0$, neglect the obstruction-probability terms in $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$, and form $\boldsymbol{\theta}^{[0]}$ using the agglomerative-clustering estimate of $\boldsymbol{\lambda}$ exactly as before. Empirical observation shows that this approach tends to provide a poor initialization for partially obstructed primitives. In particular, this approach tends to introduce large errors into the initial pose estimates of partially obstructed tophats and trihedrals. For instance, when faced with an obstructed trihedral with a response similar to that of Figure 6.2, this approach tends to produce an initial pose estimate in which the trihedral elevation is much too high. Such an estimate avoids the missed-detection penalties that would be incurred in the obstructed region by orienting much of the trihedral response above the swath of elevations that forms the 2.5° viewsphere grid (see Section 5.1.1). This estimate is often sufficiently bad that the subsequent EM iteration is unable to correct it.

The preceding observation suggests that the initialization procedure should include some mechanism for accommodating obstruction. If we could find some means of explicitly estimating $\boldsymbol{\beta}$ in the initialization, we could incorporate this estimate directly in the formation of $\boldsymbol{\theta}^{[0]}$. Explicit estimation of $\boldsymbol{\beta}$, however, would likely involve a significant increase in the computational burden of the initialization stage, due to the relatively unconstrained nature of $\boldsymbol{\beta}$ and the significant impact it can have on the appearance of primitive responses. One approach would be to iterate between estimation of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ in the initialization stage. For instance, we could form an estimate of $\boldsymbol{\theta}^{[0]}$ as in Section 4.4, use this estimate to arrive at an estimate of $\boldsymbol{\beta}$, then use this estimate of $\boldsymbol{\beta}$ to re-estimate the target parameters, and so on. This iterative refinement process could be used to obtain an initialization $\boldsymbol{\theta}^{[0]}$ that is based on an explicit estimate of $\boldsymbol{\beta}$.

We choose an approach to initialization that is simpler than that just described, but which nevertheless tends to correct much of the pose error introduced by obstruction. Our approach is a two-stage process that refines the initialization provided by the existing algorithm's initialization stage to mitigate the effects of obstruction on the pose estimate. In particular, the first stage of this process is to form an estimate of $\boldsymbol{\theta}$ exactly as in Section 4.4, *i.e.*, neglecting the presence of the $\boldsymbol{\beta}$ terms by implicitly setting them all to zero. This first stage is used to provide the type, base amplitude, location, and radius-of-curvature components of $\boldsymbol{\theta}^{[0]}$. The pose component of each primitive initialization is then obtained by maximizing over only the amplitude terms of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$, neglecting all detection-penalty terms. In other words, if the agglomerative clustering produces a report group

$$\mathcal{Z}_i = \{\mathbf{Z}_{k_1,j_1}, \ldots, \mathbf{Z}_{k_m,j_m}\} \tag{6.24}$$

for which the original initialization procedure produces a type initialization $\theta_i^{t[0]}$ and a base amplitude initialization $\theta_i^{a[0]}$, the second stage of our approach obtains the pose initialization by maximizing

$$-\frac{1}{2\sigma_a^2}\sum_{n=1}^{m}\left(a_{k_n,j_n} - A(\boldsymbol{\theta}_i, \psi_{k_n}, \phi_{k_n})\right)^2 \tag{6.25}$$

over θ_i^{p} with type and base amplitude fixed at $\theta_i^{t[0]}$ and $\theta_i^{a[0]}$, respectively. By neglecting the detection-penalty components of $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ and focusing entirely on the amplitudes of the available measurements this approach tends to produce a good pose initialization even when obstruction is present.

6.3.5 Selection of P_b

We have described an algorithm that accommodates obstruction by including a vector of hidden obstruction parameters $\boldsymbol{\beta}$ in the measurement model. Our assumed prior distribution (6.14) for $\boldsymbol{\beta}$ is a simple model that is completely specified by the selection of a scalar parameter P_b. Intuitively, P_b represents the prior probability that an arbitrary primitive is obstructed at a randomly selected viewing angle. Thus one way to choose P_b would be to select a value that is believed to approximate the "true" prior probability of obstruction.

Another way to specify P_b is suggested by examining its impact on the algorithm. Intuitively, we would expect the algorithm to be more likely to attribute a missed detection to obstruction (*i.e.*, to assign a high value for $\Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$) as P_b

increases. We can pursue this line of inquiry by rewriting the measurement model of (6.15) to make the relationship between the blockage parameters and missed detections more explicit:

$$p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta}) = \prod_{k=1}^{K} \left[p(\mathbf{Z}_k|\boldsymbol{\lambda}_k, \boldsymbol{\theta}) \cdot \frac{e^{-\gamma_{FA}V}(\gamma_{FA}V)^{F_k}}{M_k!} \cdot \prod_{i=1}^{N} l(\delta_{i,k}, \beta_{i,k}) \right], \tag{6.26}$$

where

$$l(\delta_{i,k}, \beta_{i,k}) = \begin{cases} P_D{}'_{k,i}(1 - P_b), & \delta_{i,k} = 1 \text{ and } \beta_{i,k} = 0, \\ (1 - P_D{}'_{k,i})(1 - P_b), & \delta_{i,k} = 0 \text{ and } \beta_{i,k} = 0, \\ P_b, & \delta_{i,k} = 0 \text{ and } \beta_{i,k} = 1. \end{cases} \tag{6.27}$$

(Note that a combination of $\delta_{i,k} = 1$ and $\beta_{i,k} = 1$ is impossible by assumption.) In other words, when faced with a missed detection ($\delta_{i,k} = 0$), the algorithm is forced to make a choice between two explanations ($\beta_{i,k} = 0$ or $\beta_{i,k} = 1$) that result in different values of $p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta})$. By its nature as an ML estimator, we would expect the algorithm to choose the more likely explanation, *i.e.*, the choice of $\beta_{i,k}$ that maximizes $l(\delta_{i,k}, \beta_{i,k})$ in (6.27):

$$\begin{aligned} (1 - P_D{}'_{k,i})(1 - P_b) > P_b : &\quad \text{choose } \beta_{i,k} = 0; \\ (1 - P_D{}'_{k,i})(1 - P_b) < P_b : &\quad \text{choose } \beta_{i,k} = 1. \end{aligned} \tag{6.28}$$

Thus the equation

$$(1 - P_D{}'_{k,i})(1 - P_b) = P_b \tag{6.29}$$

implicitly defines an amplitude threshold A_0 for determining whether a missed detection will be more likely to be attributed to obstruction or to simple bad luck. In particular, recalling that $P_D{}'_{k,i}$ is shorthand notation for $P_D(A(\boldsymbol{\theta}_i, \psi_k, \phi_k))$, we can express the relationship between P_b and A_0 as

$$P_b = \frac{1 - P_D(A_0)}{2 - P_D(A_0)}, \tag{6.30}$$

or, conversely,

$$A_0 = P_D^{-1}\left(\frac{1 - 2P_b}{1 - P_b}\right) \tag{6.31}$$

for any $0 \le P_b \le 1/2$, where $P_D^{-1}(\cdot)$ is the inverse of the one-to-one function $P_D(\cdot)$.

For the experiments of this chapter (and those of Chapter 7) we set $P_b = 0.25$. This corresponds to an amplitude threshold A_0 of approximately 6.0 dBsm, or equivalently to a $P_D{}'_{k,i}$ of approximately 0.67. In other words, when faced with a missed detection of primitive i at a viewing angle (ψ_k, ϕ_k) from which the primitive was expected to produce a report significantly brighter than 6.0 dBsm, the algorithm will tend to produce $\Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \approx 1$. If the predicted amplitude is significantly lower than 6.0 dBsm, we will tend to observe $\Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \approx 0$. Although $P_b = 0.25$ is a heuristic choice, we present results in Section 6.4.1 that demonstrate that the algorithm is not especially sensitive to the precise choice of P_b.

■ 6.4 Experimental Results

In this section we examine the performance of our modified algorithm on each of several targets. We begin in Section 6.4.1 by examining the modified algorithm's performance on the obstructed primitives of Section 6.1 and by examining the robustness of this performance to different choices of P_b. In Section 6.4.2 we examine the performance of the modified algorithm on three multiple-primitive targets.

■ 6.4.1 Single-Primitive Targets

Recall the partially obstructed primitives of Section 6.1. The performance of the unmodified algorithm on these primitives was summarized in Table 6.3. This performance was seen to be significantly worse than that obtained from unobstructed primitives (see Section 5.2), especially in P_{disc} and the pose and base amplitude estimate errors. Recall also that the results of Table 6.5, obtained from an algorithm that was given perfect knowledge of the obstruction, suggested that most of the degradation in the tophat and trihedral pose and base amplitude statistics was due to the mismatch between the assumed responses of the primitives and their actual obstructed responses, and not to the decrease in primitive observability caused by obstruction. (The degradation in these statistics for the dihedral and cylinder was much less marked and was seen to be approximately equally attributable to the response mismatch and to the decline in observability.)

Table 6.6 presents the performance of the modified algorithm on the same obstructed primitives. (These results correspond to 250 discoveries of each primitive at each AVSD.) To facilitate comparison to results from the old algorithm (see Table 6.3) and from the perfect-knowledge algorithm (see Table 6.5), we list the pose and base amplitude estimate error statistics for all three algorithms in Table 6.7. (All other statistics are nearly identical between algorithms.) This table demonstrates that the modified algorithm is quite successful at correcting the effects of obstruction, and generally comes close to the performance of an ideal algorithm with perfect knowledge of obstruction. For example, the modified algorithm dramatically reduces the trihedral and tophat azimuth/elevation pose estimate errors: the trihedral pose error has been reduced from 15.949° to 4.797° in the 10°-AVSD experiment (for which the perfect-knowledge value is 2.808°) and from 16.515° to 8.475° in the 20°-AVSD experiment (for which the perfect-knowledge value is 6.085°). Similarly, the tophat pose error is reduced from 12.323° to 2.984° in the 10°-AVSD experiment (for which the perfect-knowledge result is 1.621°) and from 18.518° to 4.260° in the 20°-AVSD experiment (for which the

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse [cm]
				bias	stdev	$\|\text{bias}\|$	$\|\text{stdev}\|$	az/el	rot	
10°	tri	1.000	1.000	−1.188	0.586	9.219	3.420	4.797	23.865	—
	top	1.000	1.000	−1.537	0.302	0.546	1.432	2.984	—	1.144
	dih	0.798	1.000	−2.605	3.388	5.955	16.771	17.562	16.063	—
	cyl	0.155	0.932	0.119	1.860	2.805	19.374	1.718	—	8.119
20°	tri	0.992	0.992	−1.451	1.301	9.791	6.550	8.475	30.573	—
	top	1.000	1.000	−1.490	0.511	0.487	2.743	4.260	—	2.170
	dih	0.181	0.941	−4.050	4.798	7.314	26.497	18.897	22.639	—
	cyl	0.026	0.930	0.512	2.465	2.926	33.751	4.056	—	11.497

Table 6.6. Performance of modified algorithm on partially obstructed primitives.

prim	alg	10° AVSD				20° AVSD			
		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]	
		az/el	rot	bias	stdev	az/el	rot	bias	stdev
tri	old	15.949	31.389	−1.644	1.294	16.515	32.324	−1.423	1.635
	new	4.797	23.865	−1.188	0.586	8.475	30.573	−1.451	1.301
	perfect	2.808	16.500	−0.670	0.516	6.085	27.121	−0.837	0.992
top	old	12.323	—	−3.510	0.736	18.518	—	−2.718	1.367
	new	2.984	—	−1.537	0.302	4.260	—	−1.490	0.511
	perfect	1.621	—	−1.216	0.265	3.460	—	−1.064	0.461
dih	old	18.258	20.621	−3.474	3.311	18.649	23.989	−3.926	4.978
	new	17.562	16.063	−2.605	3.388	18.897	22.639	−4.050	4.798
	perfect	15.116	15.188	−2.356	3.177	17.970	20.907	−3.673	4.971
cyl	old	2.412	—	−0.278	2.037	3.873	—	0.499	2.522
	new	1.718	—	0.119	1.860	4.056	—	0.512	2.465
	perfect	1.621	—	0.758	2.456	3.449	—	0.871	3.306

Table 6.7. Comparison of pose and base amplitude estimate statistics between algorithms for partially obstructed primitives. Here "old" refers to the algorithm described in previous chapters, "new" refers to the modified algorithm described in this chapter, and "perfect" refers to the algorithm given perfect knowledge of the obstructed regions of the viewsphere. Statistics for the old, new, and perfect algorithms are taken from Tables 6.3, 6.6, and 6.5, respectively.

perfect-knowledge result is 3.460°). The trihedral and tophat base amplitude estimates also improve with the new algorithm: for instance, in the 10°-AVSD experiment the tophat base amplitude standard deviation decreases from 0.736 dBsm to 0.302 dBsm (the perfect-knowledge result is 0.265 dBsm) and the bias changes from −3.510 dBsm to −1.537 dBsm, which is much closer to the perfect-knowledge result of −1.216 dBsm.

The new algorithm has a less dramatic effect on the dihedral and cylinder estimate statistics; as described in Section 6.1, obstruction has a less marked effect on the quality of these estimates and there is thus less room for improvement. Table 6.7 shows that

prim	P_b	10° AVSD				20° AVSD			
		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]	
		az/el	rot	bias	stdev	az/el	rot	bias	stdev
tri	0.10	4.881	26.578	−1.623	0.736	10.064	28.340	−1.822	1.421
	0.25	4.797	23.865	−1.188	0.586	8.475	30.573	−1.451	1.301
	0.40	4.784	27.236	−0.900	0.584	8.480	29.561	−1.305	1.072
top	0.10	3.618	—	−2.093	0.376	6.581	—	−2.102	0.722
	0.25	2.984	—	−1.537	0.302	4.260	—	−1.490	0.511
	0.40	2.409	—	−1.151	0.264	3.878	—	−1.178	0.466

Table 6.8. Comparison of pose and base amplitude estimate statistics for different choices of P_b.

the dihedral and cylinder azimuth/elevation pose errors have both been reduced slightly in the 10°-AVSD experiments—the dihedral error decreases from 18.258° to 17.562° (its perfect-knowledge result is 15.116°) and the cylinder error decreases from 2.412° to 1.718° (its perfect-knowledge result is 1.621°). The error statistics of the dihedral and cylinder in the 20°-AVSD experiments are nearly constant for any of the three algorithms.

To examine the robustness of the modified algorithm to different choices of P_b, we ran additional trihedral and tophat experiments for two additional values of P_b, 0.1 and 0.4. As indicated in (6.31), any value of P_b between 0 and 0.5 defines an implicit amplitude threshold A_0 at which missed detections are equally likely to be attributed to bad luck or obstruction; for $P_b = 0.1$ we have $A_0 = 7.25$ dBsm and for $P_b = 0.4$ we have $A_0 = 4.38$ dBsm. We performed 200 algorithm runs at each AVSD for both primitives using each P_b. The pose and base amplitude statistics obtained from all three choices of P_b are presented in Table 6.8. The results of this table demonstrate that the algorithm is relatively robust to the choice of P_b: performance shows some variation but is similar for each value of P_b selected. Tophat pose error decreases slightly for larger values of P_b: for instance, the 10°-AVSD tophat pose errors for P_b values of 0.10, 0.25, and 0.40 are 3.618°, 2.984°, and 2.409°, respectively. This effect is attributable to the smaller A_0 value associated with larger values of P_b, which results in more of the missed detections in the tophat's obstructed region being correctly assigned $\beta_{i,k}$ values of 1. (Changes in A_0 have a much smaller effect on the trihedral pose errors because the RCS of this primitive in most of its obstructed region is far above A_0 for any of these choices of P_b.) Another trend that is apparent in Table 6.8 is an increased positive tendency in the base amplitude bias for larger values of P_b: as A_0 increases the algorithm incurs more missed detection penalties (see (6.19)) and this will tend to result in lower base amplitude estimates. We note that for any of the P_b values tested, the results of Table 6.8 show significant improvement from the "old algorithm" results of Table 6.7.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tri	1.000	1.000	−0.433	0.422	7.604	2.406	1.941	15.204	—
	top	1.000	1.000	−0.759	0.179	0.506	1.356	1.454	—	1.106
	dih	0.964	0.996	−0.643	2.089	5.733	12.833	11.402	6.841	—
	cyl	0.404	0.990	1.137	3.236	2.220	18.969	2.619	—	6.224
20°	tri	0.998	1.000	−0.843	0.927	7.368	4.433	4.050	23.044	—
	top	1.000	1.000	−0.754	0.303	0.442	2.587	2.371	—	1.940
	dih	0.488	0.975	−2.872	4.037	6.631	23.753	15.397	14.912	—
	cyl	0.108	0.944	2.951	5.464	4.302	34.373	4.268	—	10.186

Table 6.9. Performance of modified algorithm on the target of Section 5.4.1.

■ 6.4.2 Multiple-Primitive Targets

We now examine the performance of the algorithm on three multiple-primitive targets. The first of these targets is the multiple-primitive target of Section 5.4.1 depicted in Figure 5.5. Running the algorithm on this target provides an indication of whether the added capability of the modified algorithm has any detrimental effects on its ability to estimate the parameters of unobstructed primitives. The performance of the old algorithm on this target was presented in Table 5.13; the performance of the modified algorithm is presented in Table 6.9. (These results correspond to 500 runs for each algorithm at each AVSD.) In Table 6.10 we compare the pose and base amplitude error statistics for these two algorithms. Almost all statistics are similar between algorithms. The average difference in pose errors between algorithms for the trihedral, tophat, and dihedral is 0.098°, and the average difference in the base amplitude biases among these primitives is +0.321 dBsm. Overall, the cylinder results show the greatest difference between the two algorithms: for instance, the cylinder pose error has increased from 1.209° to 2.619° in the 10°-AVSD experiment and from 1.869° to 4.268° in the 20°-AVSD experiment. Much of the variation in the cylinder statistics is attributable to the lesser statistical significance of the cylinder results due to this primitive's low P_{disc} (the cylinder results for the original algorithm are based on $500 \times 0.082 = 41$ cylinder discoveries, and those for the modified algorithm are based on $500 \times 0.108 = 54$ discoveries). Overall, Table 6.10 indicates that the performance of the modified algorithm on unobstructed primitives is similar to that of the original algorithm.

We also examined the performance of the modified algorithm on two self-obstructing multiple-primitive targets. The first of these is depicted in Figure 6.5. This primitive is comprised of a tophat, two trihedrals, and a cylinder; the cylinder obstructs the other three primitives from various viewing angles, and the cylinder and trihedrals each partially obstruct the tophat. The parameters of these components are given in Table 6.11. To gauge the ability of the modified algorithm to correct for the errors introduced by this target's self-obstruction, we performed 250 runs of each algorithm on this target at 10° and 20° AVSDs. The results from the old algorithm are presented in Table 6.12, and the results from the modified algorithm are presented in Table 6.13. The pose and

prim	alg	10° AVSD				20° AVSD			
		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]	
		az/el	rot	bias	stdev	az/el	rot	bias	stdev
trihedral	old	2.978	17.866	−0.617	0.495	4.911	25.839	−0.900	0.959
	new	1.941	15.204	−0.433	0.422	4.050	23.044	−0.843	0.927
tophat	old	1.810	—	−1.119	0.222	2.683	—	−1.026	0.374
	new	1.454	—	−0.759	0.179	2.371	—	−0.754	0.303
dihedral	old	11.337	6.934	−1.696	2.411	13.483	15.391	−2.873	4.045
	new	11.402	6.841	−0.643	2.089	15.397	14.912	−2.872	4.037
cylinder	old	1.209	—	0.154	1.507	1.869	—	0.517	2.567
	new	2.619	—	1.137	3.236	4.302	—	2.951	5.464

Table 6.10. Comparison of pose and base amplitude estimate statistics between algorithms for the target of Section 5.4.1. Here "old" refers to the algorithm described in previous chapters and "new" refers to the modified algorithm described in this chapter. Statistics for the old and new algorithms are taken from Tables 5.13 and 6.9, respectively.

primitive	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}^T}$ [cm]	θ_i^r [cm]	$\theta_i^{\mathbf{P}^T}$ [°]
tophat	2	10.00	[30.5 0.0 15.2]	18.39	[90 0 —]
trihedral 1	1	23.75	[−25.4 76.2 15.2]	—	[35.26 135 0]
trihedral 2	1	23.75	[−25.4 −76.2 15.2]	—	[35.26 −135 0]
cylinder	4	10.00	[−42.4 0 33.0]	17.70	[0 90 —]

Table 6.11. Parameters of components of first self-obstructing target.

base amplitude estimate statistics obtained from the two algorithms are directly compared in Table 6.14. We see that the modified algorithm reduces the pose errors for all primitives: for instance, the tophat pose error is reduced from 6.631° to 3.223° in the 10°-AVSD experiment and from 9.653° to 4.627° in the 20°-AVSD experiment. As with the previous experiments, we also observe an increased positive tendency in the base amplitude estimate biases and a reduction in the standard deviations of these estimates when the modified algorithm is used: for instance, in the 10°-AVSD experiment the average base amplitude estimate bias of the two trihedrals changes from −2.803 dBsm to −1.884 dBsm and the overall standard deviation of these estimates decreases from 1.106 dBsm to 0.718 dBsm.

The second self-obstructing target is depicted in Figure 6.6. This target differs from those examined thus far in that it is not a collection of discrete canonical primitives, but rather a continuous body that can presumably be approximated by several discrete primitives (*e.g.*, four trihedrals). Note that approximation of this target by such primitives will be imperfect due to the fact that components of this target are not equally sized square plates, as in the canonical trihedral. For instance, the large horizontal

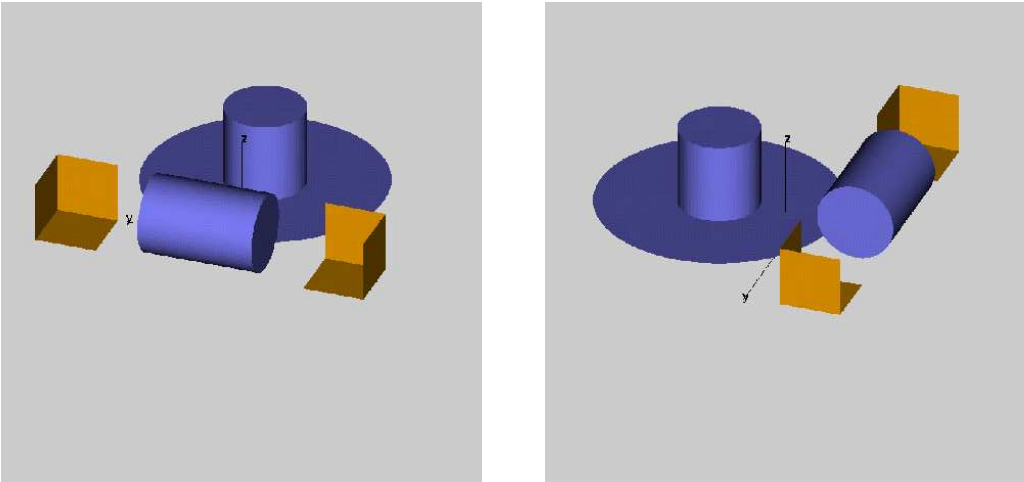


Figure 6.5. Two views of first self-obstructing target.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	top	1.000	1.000	−2.215	0.393	0.699	1.445	6.631	—	1.124
	tri 1	1.000	1.000	−2.867	1.100	6.826	2.958	6.903	24.423	—
	tri 2	1.000	1.000	−2.739	1.111	6.604	2.835	6.648	21.514	—
20°	top	1.000	1.000	−1.933	0.685	0.844	2.830	9.653	—	2.215
	tri 1	0.996	1.000	−3.267	1.841	6.482	5.653	10.548	28.075	—
	tri 2	0.996	0.996	−3.078	1.815	6.279	5.346	10.648	29.503	—

Table 6.12. Performance of old algorithm on first self-obstructing target.

		P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
				bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	top	1.000	1.000	−1.144	0.226	0.418	1.385	3.223	—	1.087
	tri 1	1.000	1.000	−1.905	0.683	6.852	2.946	5.137	22.126	—
	tri 2	1.000	1.000	−1.864	0.752	6.627	2.835	4.977	20.480	—
20°	top	1.000	1.000	−1.101	0.441	0.324	2.583	4.627	—	2.052
	tri 1	0.996	1.000	−2.729	1.445	6.447	5.574	7.362	28.601	—
	tri 2	0.996	0.992	−2.683	1.523	6.481	5.187	7.457	28.015	—

Table 6.13. Performance of modified algorithm on first self-obstructing target.

prim	alg	10° AVSD				20° AVSD			
		$\hat{\theta}_i^{\mathrm{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathrm{P}}$ rmse [°]		$\hat{\theta}_i^a$ error [dBsm]	
		az/el	rot	bias	stdev	az/el	rot	bias	stdev
top	old	6.631	—	−2.215	0.393	9.653	—	−1.933	0.685
	new	3.223	—	−1.144	0.226	4.627	—	−1.101	0.442
tri 1	old	6.903	24.453	−2.867	1.100	10.548	28.075	−3.267	1.841
	new	5.137	22.162	−1.905	0.683	7.362	28.602	−2.739	1.445
tri 2	old	6.648	21.514	−2.739	1.111	10.648	29.503	−3.077	1.815
	new	4.977	20.480	−1.864	0.752	7.457	28.015	−2.683	1.523

Table 6.14. Comparison of pose and base amplitude estimate statistics between algorithms for the first self-obstructing target. Here "old" refers to the algorithm described in previous chapters and "new" refers to the modified algorithm described in this chapter. Statistics for the old and new algorithms are taken from Tables 6.12 and 6.13, respectively.

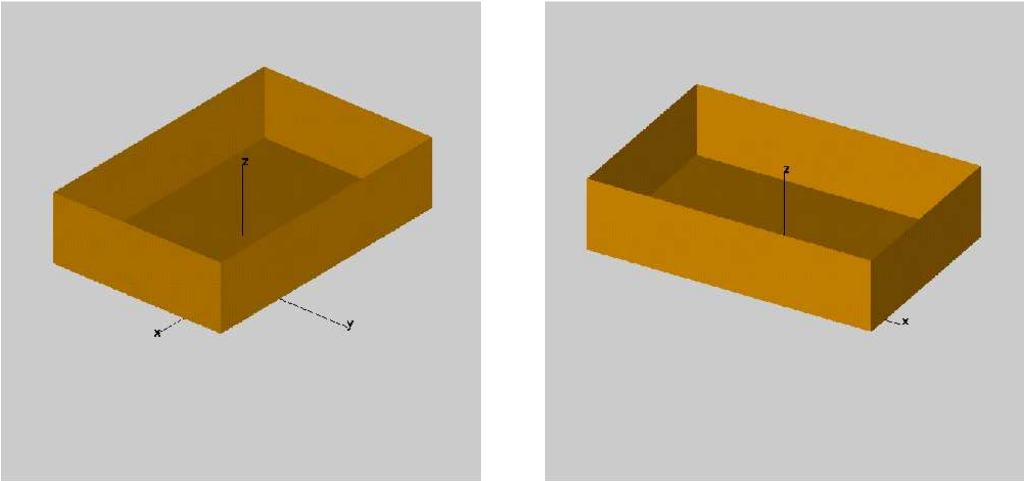


Figure 6.6. Two views of second self-obstructing target.

base of this target will result in a stronger response at unobstructed low elevations than would be observed from a canonical trihedral. These deviations from ideality can be expected to introduce additional pose errors, even on top of those introduced by obstruction. Noncanonicities such as these are the subject of Chapter 7, in which we examine approaches to detecting and accommodating noncanonical primitives. For the experiments here, our focus is simply to determine whether the modified algorithm can model this target more accurately than the original algorithm, *i.e.*, whether the modified algorithm can remove some of the errors in the pose estimate.

Table 6.15 presents parameterizations for four trihedrals that can approximate the target of Figure 6.6. (Pose is defined as for canonical trihedrals.) Because of the unequal base lengths along each dimension of this target, the canonical definition of trihedral

primitive	θ_i^t	θ_i^a [dBsm]	$\theta_i^{\mathbf{x}T}$ [cm]	θ_i^r [cm]	$\theta_i^{\mathbf{P}T}$ [°]
trihedral 1	1	> 33.3	[91.4 61.0 30.5]	—	[35.26 135 0]
trihedral 2	1	> 33.3	[−91.4 61.0 30.5]	—	[35.26 45 0]
trihedral 3	1	> 33.3	[−91.4 −61.0 30.5]	—	[35.26 −45 0]
trihedral 4	1	> 33.3	[91.4 −61.0 30.5]	—	[35.26 −135 0]

Table 6.15. Parameters of trihedral components of second self-obstructing target. Note that because of the noncanonical nature of this target, the definition of base amplitude in Table 2.1 is inapplicable; a canonical square-plate trihedral with dimensions of 45.7 cm (the shortest dimension of each of these trihedrals) has a base amplitude of 33.3 dBsm.

AVSD	P_{disc}	P_{id}	$\hat{\theta}_i^a$ stats [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
			mean	stdev	‖bias‖	‖stdev‖	az/el	rot
10°	1.000	1.000	35.195	1.920	2.576	1.287	15.908	40.484
20°	1.000	1.000	34.749	2.590	2.531	2.902	16.800	35.830

Table 6.16. Performance of old algorithm on second self-obstructing target.

base amplitude in Table 2.1 is inapplicable. (A canonical trihedral with a base length equal to 45.7 cm, the shortest dimension of each trihedral component of this target, would give a base amplitude of 33.3 dBsm.) As with the previous target, we ran two sets of experiments on the target of Figure 6.6, one using the old algorithm and one using the new algorithm. The results from 125 runs of each algorithm at each AVSD are presented in Tables 6.16 and 6.17, respectively. (These results are in ensemble form, *i.e.*, they are averages over all four trihedral components, since each corner of the target is identical except for a translation and rotation.) Comparison of these two tables reveals that the modified algorithm produces estimates with a lower azimuth/elevation pose error than the old algorithm, though with a similar rotational error. In particular, the modified algorithm reduces the azimuth/elevation pose error from 15.908° to 9.863° in the 10°-AVSD experiment and from 16.800° to 10.137° in the 20°-AVSD experiment. (The relatively large rotational error observed for these trihedrals is attributable to the fact that the best fit of the canonical unobstructed trihedral response to each of the observed obstructed responses is obtained at a rotation angle near 60°.) The modified algorithm also reduces the standard deviation of the base amplitude estimates, from 1.920 dBsm to 1.307 dBsm in the 10°-AVSD experiment and from 2.590 dBsm to 2.066 dBsm in the 20°-AVSD experiment.

AVSD	P_{disc}	P_{id}	$\hat{\theta}_i^a$ stats [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]	
			mean	stdev	‖bias‖	‖stdev‖	az/el	rot
10°	1.000	1.000	35.549	1.307	2.504	1.272	9.863	43.333
20°	1.000	1.000	35.155	2.066	2.557	2.705	10.137	39.297

Table 6.17. Performance of modified algorithm on second self-obstructing target.

alg	10° AVSD				20° AVSD			
	$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^{a}$ error [dBsm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^{a}$ error [dBsm]	
	az/el	rot	mean	stdev	az/el	rot	mean	stdev
old	15.908	40.484	35.195	1.920	16.800	35.830	34.749	2.590
new	9.863	43.333	35.549	1.307	10.137	39.297	35.155	2.066

Table 6.18. Comparison of pose and base amplitude estimate statistics between algorithms for the second self-obstructing target. Here "old" refers to the algorithm described in previous chapters and "new" refers to the modified algorithm described in this chapter. Statistics for the old and new algorithms are taken from Tables 6.16 and 6.17, respectively.

Chapter 7

Dealing with Noncanonicity

IN the previous chapter we described an extension of our basic framework to accommodate partially obstructed primitives. In this chapter we consider how our framework can be further extended to accommodate another important aspect of real targets: noncanonicity. Real targets are not comprised of ideal primitives. Instead, the motivation for reflector-primitive model-based ATR systems is the observation that many real targets can be *approximated* by collections of primitives, and that given a proper parameterization of these primitives it is possible to describe the response of a target from any viewing angle. From the perspective of model-based ATR, then, any phenomenon or target component response that cannot be adequately described in terms of the chosen parameterization represents a noncanonicity. In our framework, noncanonicities include such effects as responses from unmodeled primitives, responses from perturbed canonical primitives, and responses generated from the interaction between primitives.

In this chapter we describe an extension to our framework that is general enough to accommodate a wide range of possible noncanonicities. As described in Section 7.1, because the range of possible noncanonicities is extremely broad, our implementation focuses on one type of noncanonicity that is especially important in properly modeling a target like SLICY (see Figure 1.2) as a collection of reflector primitives. We examine SLICY in detail in Section 7.2, and demonstrate how the algorithm described in previous chapters imperfectly models SLICY's noncanonical responses. In Section 7.3 we consider various approaches to accommodating noncanonicity explicitly. In Section 7.4 we propose a particular approach that is conceptually broad enough to enable modeling of a wide range of noncanonicities without necessarily imposing a large increase in the computational burden of the algorithm. We go on to describe the implementation of this approach for a specific class of noncanonicity observed in SLICY, and describe the performance of our modified algorithm on SLICY in Section 7.5. We conclude the chapter in Section 7.6 with a description of a further modification to our algorithm that enables identification of target features that are improperly modeled using the specified primitive set and parameterization.

7.1 Examples of Noncanonicity

The range of possible noncanonicities is extremely broad. Given any set of primitives (*e.g.*, the trihedral, tophat, dihedral, and cylinder of Figure 2.10), a parametric framework for describing a target comprised of such primitives (*e.g.*, $\boldsymbol{\theta}$ as described in Section 3.1), and a measurement model relating the target parameters to the observed features in a set of SAR images (*e.g.*, $p(\boldsymbol{\lambda}, \boldsymbol{\beta}, \mathbf{Z}|\boldsymbol{\theta})$ as constructed in Section 3.3), a noncanonicity is any observed response that is not generated strictly according to the measurement model, or that involves primitive attributes not captured by the chosen parameterization, or that is generated by an unmodeled response mechanism. In this section we describe several basic sources of noncanonicity in target responses. In Section 7.3 we consider how various types of noncanonicity could be accommodated by augmentation of the parameterization $\boldsymbol{\theta}$ to include additional descriptive parameters, modification to the measurement model, or a combination of these approaches.

One type of noncanonicity corresponds to responses from unmodeled primitives. For instance, returns from flat plates, cavities, edges, corners, and other such target components are noncanonical in our existing framework because we have chosen to model only trihedrals, tophats, dihedrals, and cylinders. Realistic modeling of targets whose responses include many unmodeled scattering mechanisms would require augmentation of the set of primitives under consideration and specification of a scattering-response function $S_{\theta_i^t}(\cdot)$ for any additional primitives we wish to model.

Another type of noncanonicity arises when primitives are perturbed from their canonical geometric forms. This perturbation can take several forms. For primitives comprised of several components, changes in the relationship between these components will lead to noncanonical responses. For instance, small deviations in the internal angle of a dihedral can have a significant impact on the primitive's angular scattering response [3, 39, 60]; trihedral responses tend to be more robust to small deviations in internal angle but also display a dependence on such variations [21,33]. Modifications in the size or shape of one or more components of a primitive will also lead to noncanonicity. For instance, a trihedral response can be generated by three flat plates even if they are not square plates; the overall response of such a trihedral, however, will differ from the canonical square-plate trihedral response described by $S_1(\cdot)$.[1] Similarly, a tophat formed from a cylinder shaft and a noncircular base will not have a radially symmetric response as predicted by $S_2(\cdot)$, but will generate stronger returns at azimuths from which its base extent is greater.

The types of noncanonicity just described conceptually lend themselves to compact representation. Given a handful of parameters describing the plate sizes and internal angle of a dihedral, for instance, it is possible at least in theory to predict the response of such a primitive using PO or similar means. Other types of noncanonicity do not lend themselves to such compact representation in terms of a small number of primitive

[1] A detailed examination of the responses of various trihedrals formed from non-square plates is presented in [94].

attributes. For instance, the presence of small dents or scratches on the surface of a primitive will alter its response, but would be much harder to characterize compactly.

Some types of noncanonicity are essentially stochastic in nature. For instance, surface roughness will lead to variations in primitive response that generally cannot be deterministically modeled with much accuracy [62,92,104]. Similarly, partial or uneven obscuration by dielectric materials such as foliage, mud, or camouflage will all alter the observed response of a primitive and are often difficult to describe deterministically or in terms of a handful of parameters [62,104].

These examples illustrate that the range of potential noncanonicities is enormous. Rather than attempting to deal with each possible type of noncanonicity explicitly, our goal in this chapter is to construct a framework that is flexible enough to allow significant freedom in the choice of *which* types of noncanonicity to model explicitly for a given implementation of the target model generation algorithm, and to demonstrate this framework by modeling a particular class of noncanonicity that is especially important in describing the responses observed from SLICY.

■ 7.2 SLICY

Figure 7.1 depicts four views of an XPatch facetization model based on SLICY. As described in Section 1.2, SLICY is a benchmark target that exhibits many sophisticated real-world effects, as described shortly. Examination of Figure 7.1 suggests that SLICY might be effectively modeled as a collection of reflector primitives. In fact, SLICY contains components that are similar to each of our four basic primitive types. We detail these components below. To aid in the description of SLICY's components, Figure 7.2 depicts a view of SLICY with cardinal azimuth directions (defined as depicted in Figure 2.9) clearly indicated. We define the front of SLICY to be the side clearly visible at 90° azimuth, as indicated in Figure 7.2.

- *Trihedrals.* SLICY contains three components that can be modeled as trihedrals. Two of these are cut out from the front of SLICY's rectangular body; the third is formed by two triangular plates that extend from the top of SLICY. Note that none of these trihedrals is a canonical square-plate trihedral as depicted in Figure 2.10: the two "cut-out" trihedrals are each formed from two rectangular plates and a quarter-circular plate, and the "pop-up" trihedral is formed from two triangular plates and an extended segment of SLICY's large top.

- *Tophats.* The two cylindrical shafts extending from the top of SLICY can be modeled as tophats. The short tophat on the left is hollow; the tall tophat on the right is solid. Like the trihedrals just described, SLICY's tophats are noncanonical: their bases are formed not from circular plates as in Figure 2.10, but from SLICY's large rectangular top.

- *Dihedrals.* Although SLICY contains no true dihedral components that are distinct from its trihedrals, all three trihedrals generate dihedral-like double-bounce returns from some viewing angles. The two cut-out trihedrals on the front of SLICY will obviously produce dihedral returns when viewed at azimuths near 90°; these components are large enough that the dihedral responses will generally be resolvable from the cut-out trihedrals' triple-bounce returns. Additionally, the triangular plates that form SLICY's pop-up trihedral are visible from some viewing angles from the back of SLICY and generate double-bounce returns near azimuths $-45°$ and $-135°$. Although the physical structures generating these returns are actually trihedrals, the separation of the just-described dihedral responses from the corresponding trihedral responses in location (for the cut-out dihedral double-bounce returns) or in pose (for the pop-up dihedral double-bounce returns) often lead to these components being modeled as distinct scatterers by the target model generation algorithm, as we discuss in Sections 7.2.2 and 7.5. Although it is possible to imagine employing geometrically based reasoning to prevent the inclusion of distinct primitive estimates for responses generated by already-modeled physical structures, our algorithm contains no means for doing this. We thus include these four dihedrals in the list of SLICY's component primitives. We point out that although the two pop-up dihedrals produce returns over a different range of viewing angles than does the triangular-plate trihedral, their close proximity to that component presents a particular challenge to the estimation of that trihedral's parameters, as will be discussed further in Sections 7.2.2 and 7.5.

- *Cylinders.* SLICY contains one component that can be modeled as a cylinder. In particular, the front of SLICY contains a curved surface of constant radius-of-curvature that will produce visible returns at azimuths near 90°.

Figure 7.3 depicts two views of the SLICY facetization model that have been annotated to indicate the components just described. These components have been named and numbered for unique identification. In particular, "tophat 1" is the short hollow tophat on the left, "tophat 2" is the tall solid tophat on the right, "trihedral 1" is the long extended cut-out trihedral on the left, "trihedral 2" is the shorter (but still extended) cut-out trihedral on the right, and "trihedral 3" is the triangular-plate pop-up trihedral on top of SLICY. "Dihedral 1" is the extended cut-out dihedral on the left, "dihedral 2" is the cut-out dihedral on the right, "dihedral 3" is the dihedral formed from the SLICY's top base and the left triangular plate of the pop-up trihedral, and "dihedral 4" is the dihedral formed from the top of SLICY and the triangular plate on the right.

The parameterizations of these components of SLICY are given in Table 7.1. Pose is defined in Table 7.1 as for canonical primitives: for instance, trihedral azimuth and elevation are defined as the viewing direction equidistant from the three trihedral axes, even though noncanonical trihedrals will generally not produce their maximum responses from this viewing angle. Similarly, tophat pose is defined as the direction of the axis of the tophat cylinder shaft. Note that base amplitude is not given in Table 7.1. This is because the noncanonicity of these components makes the canonical definition

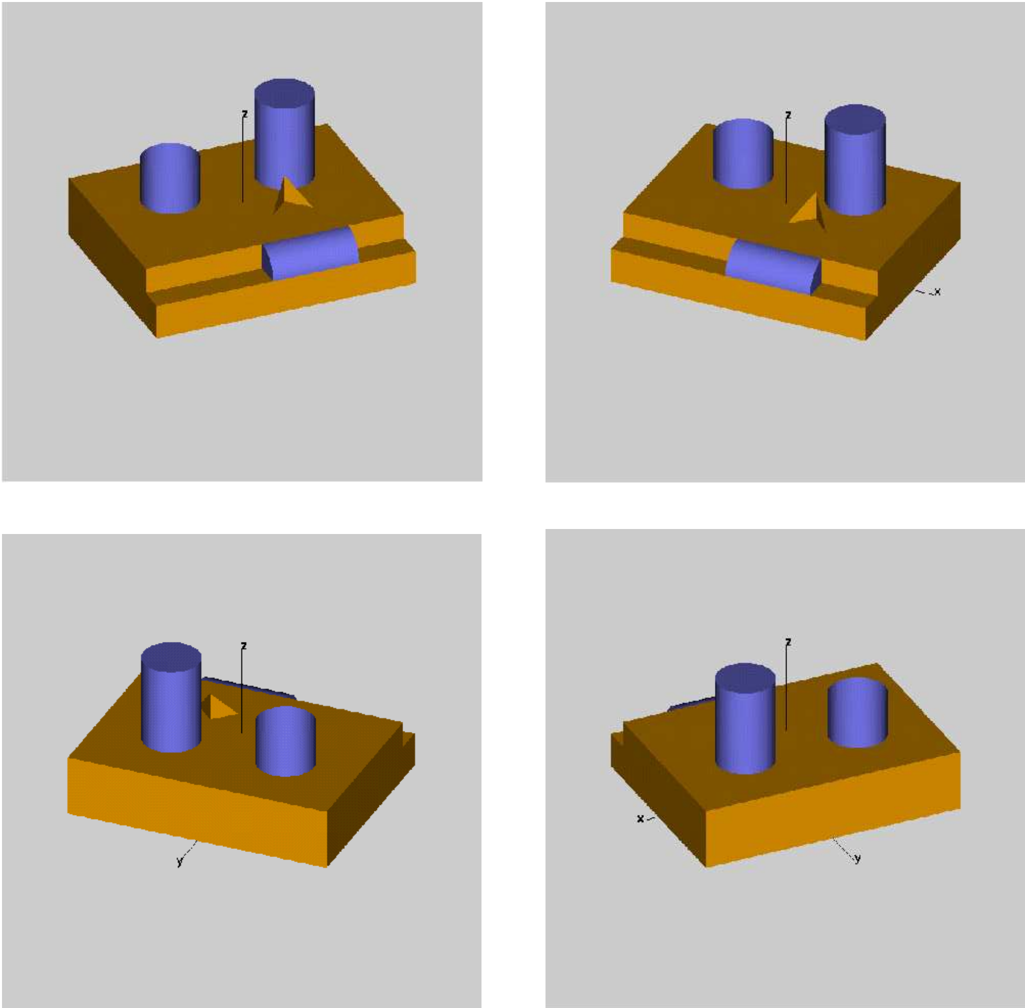


Figure 7.1. Four views of a facetization model of SLICY. This target extends 305 cm in the x direction and 244 cm in the y direction.

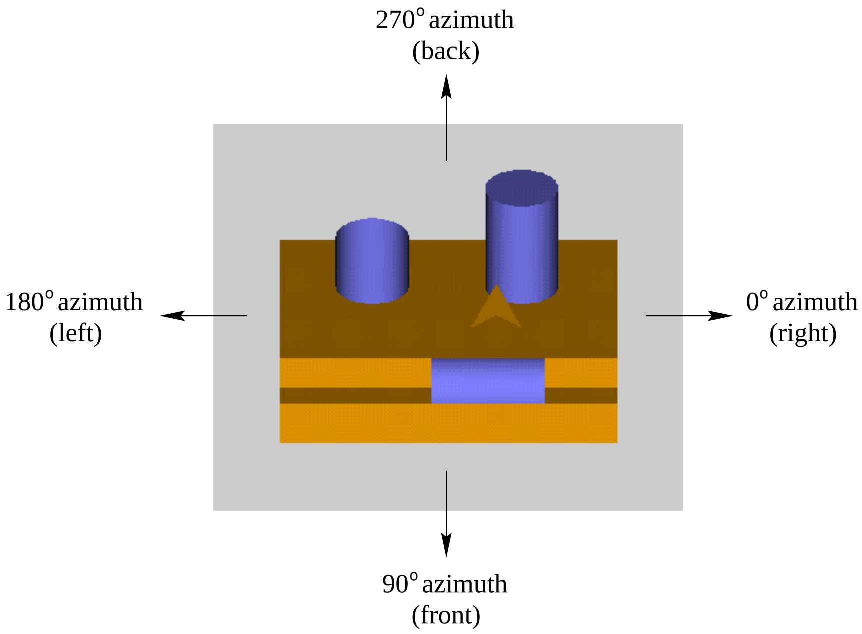


Figure 7.2. SLICY frame of reference. Here azimuth is defined as in Figure 2.9. We define the front of SLICY to be the side clearly visible at 90° azimuth. The direction of 0° azimuth corresponds to the x-axis in Figure 7.1; the direction of 270° (or, equivalently, $-90°$) azimuth corresponds to the y-axis.

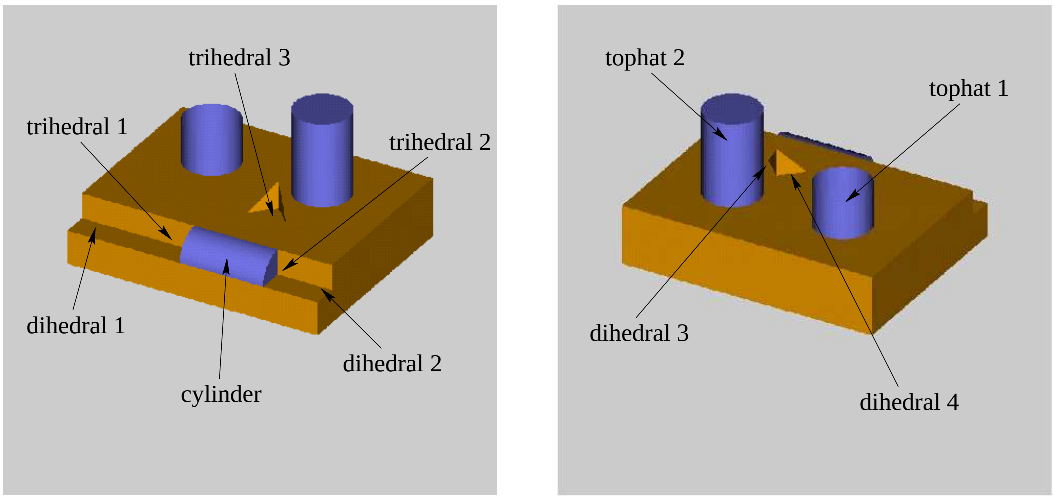


Figure 7.3. Reflector-primitive components of SLICY.

component	θ_i^t	$\theta_i^{\mathbf{x}T}$ [cm]			θ_i^r [cm]	$\theta_i^{\mathbf{p}T}$ [°]		
tophat 1	2	[−68.6	38.1	71.1]	33.0	[90	0	—]
tophat 2	2	[66.0	38.1	71.1]	33.0	[90	0	—]
trihedral 1	1	[−15.2	−91.4	40.6]	—	[35.26	135	0]
trihedral 2	1	[86.4	−91.4	40.6]	—	[35.26	45	0]
trihedral 3	1	[43.2	−15.2	71.1]	—	[35.26	90	0]
cylinder	4	[35.6	−91.4	40.6]	30.5	[0	0	—]
dihedral 1	3	[−83.8	−91.4	40.6]	—	[45	90	90]
dihedral 2	3	[119.4	−91.4	40.6]	—	[45	90	90]
dihedral 3	3	[31.5	−26.9	71.1]	—	[45	−45	90]
dihedral 4	3	[54.9	−26.9	71.1]	—	[45	−135	90]

Table 7.1. Parameters of SLICY components.

of base amplitude (*i.e.*, the log-scaling factor that must be applied to $S_{\theta_i^t}(\cdot)$ to match the actual response of the primitive) inapplicable.

In addition to displaying responses from each of the components just described, a variety of other effects are important in understanding the responses produced by SLICY. These are as follows:

- *Obstruction.* Some of SLICY's components partially obstruct other components. In particular, tophats 1 and 2 are mutually self-obstructing along SLICY's long axis (*i.e.*, from azimuths near 0° and 180°). Trihedral 3 also partially obstructs each tophat over a region of the viewsphere. Dihedral 4 is almost completely obstructed by tophat 1 except at high elevation angles. The curved surface on the front of SLICY is not visible from the back, and thus SLICY's cylinder component is effectively obstructed over half its response.

- *Noncanonical responses from modeled primitives.* As described previously, most of the components described in Table 7.1 are noncanonical. That is, their geometries differ from the canonical primitive geometries depicted in Figure 2.10, and thus their observed angular responses will differ from those predicted by PO and GO in Table 2.1. For instance, each of SLICY's two cut-out trihedrals will produce its strongest response not at a viewing angle equidistant from its three axes, but rather at a viewing angle closer to its longest axis, since such a viewing angle will maximize the projected triple-bounce reflection area of the primitive [94]. Similarly, SLICY's tophats will produce responses that vary with azimuth. The strongest tophat responses will be produced in directions from which the projected double-bounce area of the tophat is maximized, *i.e.*, the viewing directions from which the rectangular base-plate length is largest. For SLICY's tophats this will occur near azimuths ±45° and ±135°, *i.e.*, at directions extending from each tophat toward the corners of SLICY's extended base.

- *Responses from unmodeled primitives.* As indicated previously, SLICY's short tophat is hollow. When viewed from high elevation angles, there is a clear line-of-sight to SLICY's flat top base through this cylinder shaft and a double-bounce response is produced from the interior of this tophat.[2] At lower elevation angles, this hollow tophat produces a cavity response [62, 66]. In particular, rays that enter the concave region formed by the hollow tophat are scattered multiple times before they escape the cavity; some of these rays are reflected back in the direction of the sensor. Because of the additional distance these rays travel before returning to the sensor, the cavity response appears downrange from the projected location of the tophat. Additionally, due to the dispersive nature of multiple reflections from a curved surface, such a response will generally be much dimmer than the corresponding tophat or high-elevation interior double-bounce response.

- *Multiple-primitive responses.* The proximity of the components on top of SLICY results in multiple-primitive responses (*i.e.*, responses produced by rays that are reflected from more than one primitive before returning to the sensor) at many viewing angles. In particular, tophats 1 and 2 produce strong multiple-primitive reflections over much of the viewsphere. Such responses appear downrange from the true locations of the tophats and at cross-range locations approximately midway between the two primitives. Similarly, trihedral 3 and tophat 2 produce multiple-primitive reflections at viewing angles near azimuths 0° and 180°. Due to the proximity of this pair of primitives, their multiple-primitive response often appears to emanate from a point very close to the apex of trihedral 3. As described further in Sections 7.2.2 and 7.5, this has important ramifications for the estimation of that primitive's parameters.

- *Unresolvable primitives.* The proximity of some of SLICY's components complicates the feature extraction process (see Section 3.2). In particular, from some viewing angles neighboring components are separated by less than the width of a resolution cell (*i.e.*, 0.30 m) and are unresolvable. For instance, at their closest points, the curved surface of tophat 2 and the apex of trihedral 3 are separated by only 0.25 m. From many viewing angles at which both primitives produce strong responses, the individual responses blend into a single peak. The feature extractor described in Section 3.2 thus extracts a single report for these two primitives at such viewing angles. Furthermore, the features of this report (*i.e.*, its location, amplitude, and polarimetric-signature classification) are generally a poor match to the features expected to be generated by either primitive in isolation, due to the coherent interference between the two responses. In addition to this pair of

[2] The hollow tophat has a diameter of 66.0 cm and a height of 50.8 cm; the elevation angle at which SLICY's top becomes visible through the cylinder shaft is thus arctan(50.8/66.0), or approximately 37.6°.

primitives, two other components of SLICY are sometimes unresolvable: the exterior and interior responses from tophat 1 (*i.e.*, its standard tophat double-bounce response and its unmodeled interior double-bounce response) often interfere to produce only a single peak.

SLICY thus displays a variety of sophisticated effects, only one of which (obstruction) is explicitly accommodated by our existing algorithm. In Section 7.2.2 we describe how each of these effects influences target modeling by presenting the performance of a slightly modified version of the existing algorithm. This modification is not motivated by any of the types of noncanonicity just described, but instead by a difficulty presented the existing algorithm by the size of SLICY's tophats, as we now describe.

7.2.1 A Modification to the Initialization Procedure

Recall that the initialization procedure of the existing algorithm, described in detail in Section 4.4, clusters reports under the assumption that reports generated by a single primitive all project from a single point in $\mathbb{R}^3$. In other words, the existing initialization procedure assumes that for any report $\mathbf{Z}_{k,j}$ for which $\lambda_{k,j} = i$, we have

$$\mathbf{x}_{k,j} \sim \mathbf{H}_k \theta_i^{\mathbf{x}}, \tag{7.1}$$

where $\mathbf{H}_k$ is the 2×3 ground-to-slant-plane transformation matrix for image k defined in (3.4). As indicated previously, the model of (7.1) is inaccurate for tophats and cylinders: these primitives produce reports whose locations project from a specular point $\mathbb{R}^3$ that depends on the viewing angle and on their radii:

$$\mathbf{x}_{k,j} \sim \mathbf{H}_k \theta_i^{\mathbf{x}} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \theta_i^r \cos \psi_k'(\theta_i^{\mathbf{P}}), \tag{7.2}$$

where $\psi_k'(\theta_i^{\mathbf{P}})$ is the pose-dependent relative viewing elevation of primitive i in image k (see Figure 2.10). Thus there is a mismatch between the assumed and actual report locations of tophats and cylinders. We have seen that for tophats and cylinders of the size considered in previous chapters, the assumption of (7.1) is, while imperfect, still sufficient to produce a good initialization. For tophats of size comparable to SLICY's, the mismatch between (7.1) and (7.2) is severe enough to have a significant effect on the algorithm. In particular, when the existing initialization procedure is faced with one of SLICY's tophats, it tends to produce several small clusters of reports attributable to the primitive instead of one large cluster. This is undesirable because many of these distinct initializations often survive to the end of the EM iteration, producing a target model that contains several primitives where there should be only one.

In order to cluster tophat reports properly according to (7.2), we would require knowledge of $\theta_i^{\mathbf{P}}$ at the time of agglomeration. Although we can imagine modifying the existing agglomeration procedure to cluster reports based not only on estimates of primitive location, but also on estimates of pose and radius, such a modification would vastly complicate the initialization stage. In particular, it would require primitive pose

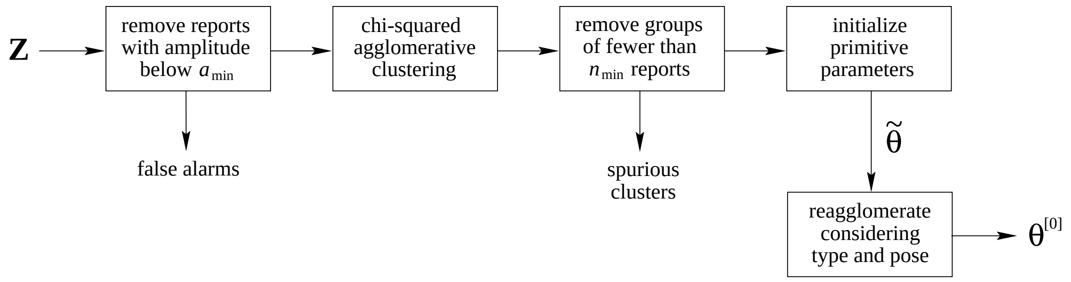


Figure 7.4. Block diagram of modified initialization procedure.

to be estimated for each pair of report groups at each agglomeration iteration. As described in Section 4.3, pose estimation is much more difficult and computationally demanding than location estimation. We thus propose to incorporate knowledge of primitive pose as a post-processing step that takes place at the conclusion of the existing initialization procedure, when only a handful of report clusters remain. This post-processing step takes the form of a second agglomeration stage in which primitive pose is considered in order to determine whether any clusters produced by the existing procedure should be further grouped. A block diagram of our modified initialization procedure, including this post-processing step, is depicted in Figure 7.4.

Let us denote the primitive parameter-vector estimates produced by the existing initialization procedure by $\tilde{\boldsymbol{\theta}}_i$, for $i = 1, \ldots, \tilde{N}$.[3] Each of these estimates is based on a cluster of reports

$$\boldsymbol{\mathcal{Z}}_i = \{\mathbf{Z}_{k_1,j_1}, \ldots, \mathbf{Z}_{k_{n_i},j_{n_i}}\}. \tag{7.3}$$

Without loss of generality let us assume that

$$n_1 \geq n_2 \geq \cdots n_{\tilde{N}}. \tag{7.4}$$

Our second agglomeration clustering stage is based on a dissimilarity measure between pairs of clusters that is very similar to that of Section 4.4, but in which (7.2) is used to form location and radius estimates for tophats and cylinders. In particular, for any two clusters $\boldsymbol{\mathcal{Z}}_{i_1}$ and $\boldsymbol{\mathcal{Z}}_{i_2}$ with $i_1 < i_2$ (so that $n_{i_1} \geq n_{i_2}$) our dissimilarity measure is the chi-squared cdf evaluated at the total squared error of the estimate associated with

[3]In the existing initialization procedure, $\boldsymbol{\theta}_i^{[0]} = \tilde{\boldsymbol{\theta}}_i$ and $N = \tilde{N}$.

those clusters, *i.e.*,

$$\tilde{\varepsilon}^2_{\mathrm{L}}(\boldsymbol{\mathcal{Z}}_{i_1}, \boldsymbol{\mathcal{Z}}_{i_2}) = \begin{cases} \displaystyle\sum_{\substack{\mathbf{Z}_{k,j}\in \\ \boldsymbol{\mathcal{Z}}_{i_1}\cup\boldsymbol{\mathcal{Z}}_{i_2}}} \left\| \mathbf{x}_{k,j} - \left(\mathbf{H}_k \hat{\mathbf{X}} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \hat{r} \cos \psi'_k(\tilde{\theta}^{\mathrm{P}}_{i_1}) \right) \right\|_{\mathbf{R}^{-1}}, & \tilde{\theta}^t_{i_1} \in \{2, 4\}, \\ \displaystyle\sum_{\substack{\mathbf{Z}_{k,j}\in \\ \boldsymbol{\mathcal{Z}}_{i_1}\cup\boldsymbol{\mathcal{Z}}_{i_2}}} \left\| \mathbf{x}_{k,j} - \mathbf{H}_k \hat{\mathbf{X}} \right\|_{\mathbf{R}^{-1}}, & \tilde{\theta}^t_{i_1} \in \{1, 3\}, \end{cases} \tag{7.5}$$

and where $\hat{\mathbf{X}}$ and $\hat{r}$ are LLSE location and radius estimates formed according to (7.2) if $\tilde{\theta}^t_{i_1}$ indicates a tophat or cylinder, or where $\hat{\mathbf{X}}$ is the LLSE location estimate formed according to (7.1) if a trihedral or dihedral is indicated.

Given an $\tilde{N} \times \tilde{N}$ matrix of dissimilarities calculated as above, the second agglomeration stage proceeds almost identically to the first. It uses the same dissimilarity threshold η. At each agglomeration iteration, if the two most similar clusters are less dissimilar than η, they are merged into a single cluster. The location and, if applicable, radius estimates for this cluster are available from the calculation of (7.5); other primitive attributes are estimated as described in Section 4.4. After $\tilde{N}$ is decremented and a new set of dissimilarity scores is calculated using the attributes of the new merged cluster, the agglomeration proceeds to the next iteration. This is continued until all remaining clusters are more dissimilar than η, at which point the remaining estimates $\tilde{\boldsymbol{\theta}}$ are passed to the EM iteration as $\boldsymbol{\theta}^{[0]}$.

This modified initialization procedure—different from the original approach only in its inclusion of the post-processing step just described—greatly reduces the tendency of the original approach to produce numerous primitive initializations for a single large-radius tophat. At the conclusion of the modified initialization procedure, there is generally a single primitive initialization for each tophat component.

■ 7.2.2 Performance of the Existing Algorithm on SLICY

We now examine the performance of the existing algorithm (*i.e.*, the algorithm described in previous chapters and whose initialization stage has been modified as described in the previous section) on SLICY. We performed 150 algorithm runs at 10° and 20° AVSDs, using the same algorithmic and measurement-model parameters that were used in previous chapters. The prior-obstruction-probability parameter P_b was set to 0.25 as in most experiments of Chapter 6.

In every run at both AVSDs, the algorithm discovered the presence of SLICY's dominant scatterers, *i.e.*, its two tophats and three trihedrals. The less observable components (*i.e.*, the dihedrals and cylinder described in Section 7.2 and parameterized in Table 7.1) were discovered in fewer runs. Table 7.2 lists the fraction of runs in which each component was discovered. Among the dihedral and cylinder components, dihedral 3 was discovered in the most runs (41% of the 10°-AVSD runs and 5% of

component	P_{disc}	
	10° AVSD	20° AVSD
tophat 1	1.00	1.00
tophat 2	1.00	1.00
trihedral 1	1.00	1.00
trihedral 2	1.00	1.00
trihedral 3	1.00	1.00
dihedral 1	0.13	0.03
dihedral 2	0.24	0.04
dihedral 3	0.41	0.05
dihedral 4	0.02	0.01
cylinder	0.05	0.02

Table 7.2. Fraction of runs in which each component of SLICY was discovered.

the 20°-AVSD runs). Dihedrals 1 and 2 have very narrow specular responses and are discovered in fewer runs. Dihedral 4 is obstructed by tophat 1 over much of its response region, and thus is discovered in very few runs. Similarly, the narrow response of the cylinder, coupled with its effective obstruction from the rear of SLICY, results in very few discoveries of this component.

The error statistics for the tophat and trihedral parameter estimates are displayed in Table 7.3. The pose-error statistics of Table 7.3 are of particular note because they demonstrate the effects of noncanonical responses from modeled primitives. For instance, although the trihedral and tophat components of SLICY are highly observable and are as bright or brighter than similar components examined in previous chapters, the pose errors of these components are larger than those observed in previous chapters: in the 10°-AVSD experiment the tophat pose errors are 4.88° and 13.22° (compare to 1.33° for the tophat of Section 5.2), and the trihedral azimuth/elevation pose errors are all greater than 8.99° (compare to 2.17° for the trihedral of Section 5.2). Figure 7.5 presents scatter plots of the azimuth and elevation estimates for the two tophats at each AVSD; Figure 7.6 presents similar scatter plots for the three trihedrals. (The range of azimuths and elevations displayed in the tophat plots is different from that in the trihedral plots in order to display the variation in pose estimates over the different portions of the viewsphere near each primitive's true pose.) Note that in each plot, the estimates are not centered around the true primitive pose, but instead around an incorrect value. For instance, the trihedral 1 estimates do not cluster around elevation 35.26° and azimuth 135°, but instead around a lower-elevation and higher-azimuth pose. This represents a reorientation of the assumed canonical trihedral response to better match the asymmetric noncanonical response actually observed from trihedral 1. Similarly, the pose estimates of tophat 2 do not cluster around elevation 90°, but instead around a value with elevation near 77° and azimuth near 135°. By orienting the assumed canonical tophat response at this angle the algorithm can better fit the

AVSD	component	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
		mean	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tophat 1	16.65	0.32	9.57	2.70	4.88	—	3.91
	tophat 2	18.98	0.80	2.35	1.34	13.22	—	0.92
	trihedral 1	27.01	0.80	2.44	1.64	10.44	30.05	—
	trihedral 2	26.53	0.66	5.83	3.07	8.99	23.33	—
	trihedral 3	25.07	1.57	8.28	3.15	13.17	26.50	—
20°	tophat 1	16.54	0.47	10.65	5.56	4.73	—	4.12
	tophat 2	19.08	1.10	2.30	2.61	13.48	—	1.56
	trihedral 1	26.40	1.43	1.83	3.17	10.85	36.71	—
	trihedral 2	25.97	1.11	5.74	6.24	10.05	29.14	—
	trihedral 3	25.68	4.69	6.87	8.13	21.95	29.13	—

Table 7.3. Results of existing algorithm on SLICY.

asymmetric noncanonical response actually observed from tophat 2, which produces its strongest return near azimuth 135° due to SLICY's large base extent in this direction. Figures 7.5 and 7.6 clearly show that noncanonicity is introducing systematic biases into primitive pose estimates, and demonstrate the need for introducing some means to describe noncanonical responses from modeled primitives. This is the focus of Sections 7.3 and 7.4.

Table 7.3 indicates that the estimates produced for trihedral 3 have a larger pose error than those produced for the other two trihedrals. For instance, in the 20°-AVSD experiment, trihedrals 1 and 2 display azimuth/elevation errors of 10.85° and 10.05°, respectively, while the corresponding error for trihedral 3 is 21.95°. This is attributable to two factors first discussed in Section 7.2. First is the unresolvability of trihedral 3 and tophat 2 from many viewing angles, and the consequential availability of only a single report corresponding to the pair of primitives at these viewing angles. This effectively reduces the observability of both primitives, but has a much more marked effect on trihedral 3 because trihedrals are fundamentally less observable than tophats. Compounding this decrease in observability is the fact that the features of a single unresolved report will generally be a poor match to those expected to be produced by either the trihedral or the tophat. Thus even when such a report is associated with trihedral 3, it will tend to introduce additional error into the estimate produced for that primitive.

The second factor leading to the larger pose error of trihedral 3 is the proximity of this trihedral's response to several other response mechanisms that produce reports at different viewing angles. In particular, dihedrals 3 and 4 can produce reports that project from locations near the apex of trihedral 3 at azimuths near −45° and −135°, respectively. Similarly, trihedral 3 and tophat 2 interact to produce multiple-primitive responses at azimuths near 0° and 180°; due to the proximity of trihedral 3 and tophat 2 these responses appear to emanate from a location near the apex of trihedral 3. The proximity of trihedral 3 to these other response mechanisms presents a particular dif-

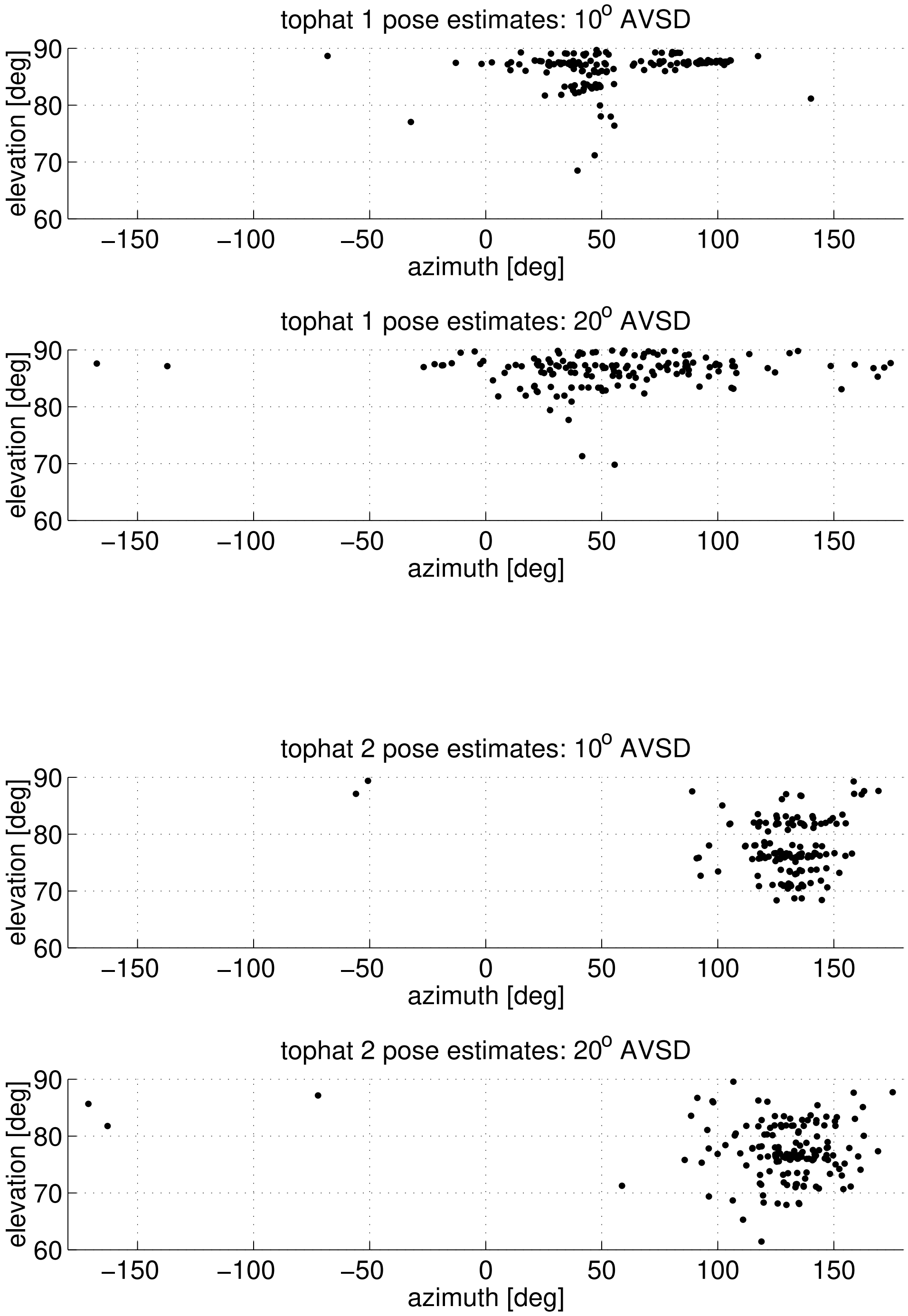


Figure 7.5. Tophat pose estimates produced by the original algorithm. The true pose of each tophat is 90° elevation, the horizontal line at the top of each plot. Notice that due to noncanonicity the pose estimates cluster around incorrect values.

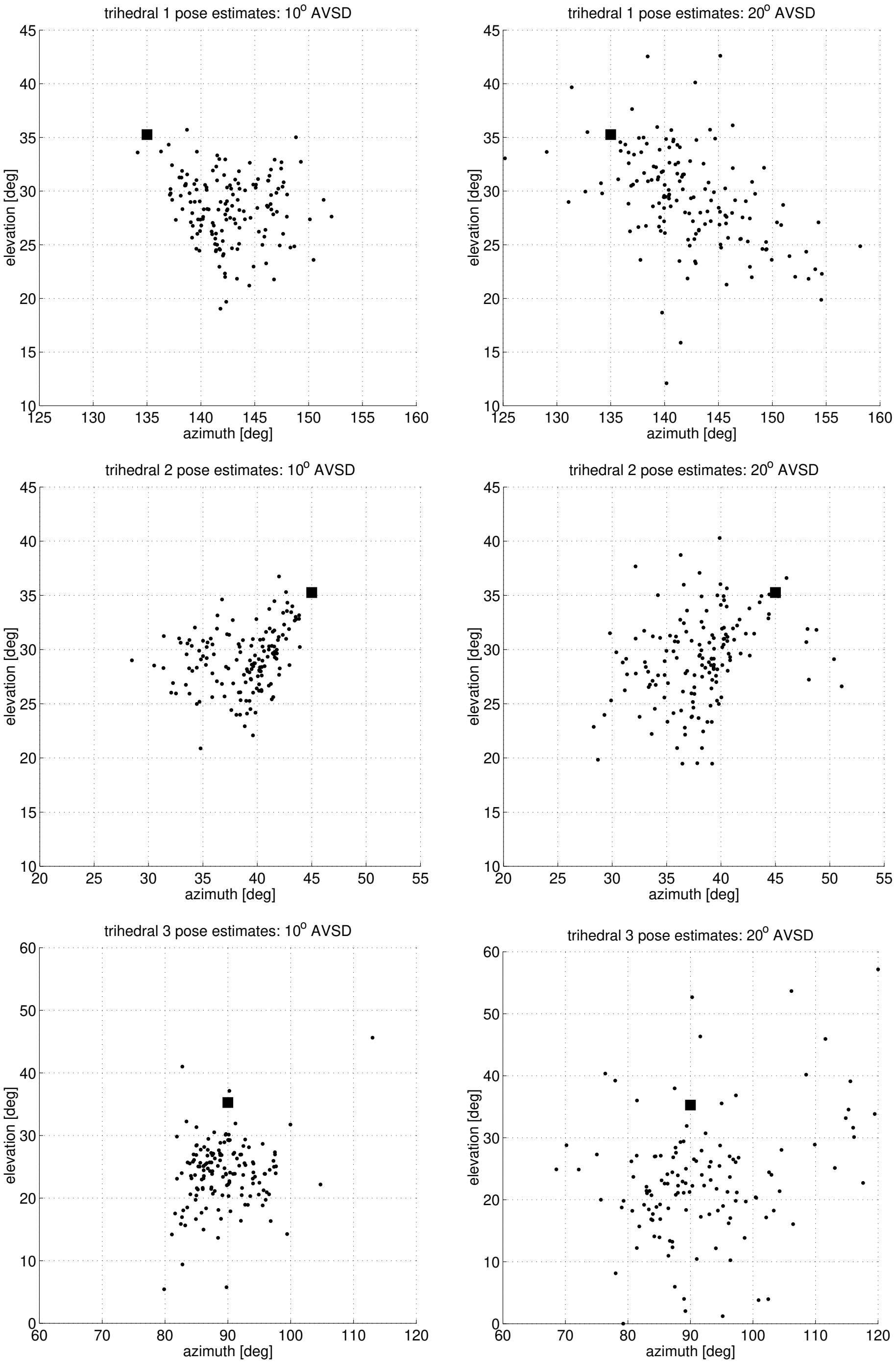


Figure 7.6. Trihedral pose estimates produced by the original algorithm. The true pose of each trihedral is denoted by the filled square in each plot. Notice that due to noncanonicity the pose estimates cluster around incorrect values.

ficulty to the algorithm: some of the reports produced by these additional response mechanisms are often grouped with the legitimate trihedral 3 reports either in the initialization stage or as the iteration progresses, leading to additional errors.

Primitive unresolvability also has an effect on the estimates produced for tophat 1. In particular, this primitive's unresolvability from its internal hollow-tophat double-bounce response at high elevation angles leads to location and radius-of-curvature errors that are much higher than those observed from tophat 2 (*e.g.*, tophat 1 exhibits a location-estimate bias of 9.57 cm in the 10°-AVSD experiment as compared to a bias of 2.35 cm for tophat 2).

In addition to the components indicated in Table 7.2, most algorithm runs resulted in at least one spurious primitive estimate being produced. In particular, in the 10°-AVSD experiment an average of 2.23 spurious estimates were produced in each run; in the 20°-AVSD experiment an average of 1.90 spurious estimates were produced in each run. Most of these spurious primitives represent attempts by the algorithm to model responses generated by multiple-primitive reflections or by the interior double-bounce response of the hollow cavity of tophat 1. Most such spurious primitives are classified as trihedrals, generally because the responses generating spurious estimates are visible over a portion of the viewsphere that is limited in extent but not so limited as to enable accurate modeling by a dihedral or cylinder. In Section 7.6 we describe a method for distinguishing between spurious primitive estimates and legitimate primitive estimates corresponding to explicitly modeled response mechanisms.

Incidentally, although noncanonicity introduces significant errors into the pose estimates obtained for SLICY's tophats, these errors are smaller than they would be if our algorithm included no means for accommodating obstruction. In particular, as previously indicated, SLICY's tophats are mutually self-obstructing at azimuths near 0° and 180°. To demonstrate that the algorithm is detecting and accommodating this obstruction, we performed 50 algorithm runs using 10°- and 20°-AVSD data sets with P_b set to 0 (*i.e.*, with the obstruction capability of the algorithm completely suppressed). The tophat pose and base amplitude statistics obtained with $P_b = 0$ are compared to those of Table 7.3 (obtained with $P_b = 0.25$) in Table 7.4. Note that with obstruction-modeling enabled, the algorithm produces tophat pose estimates with lower pose errors at each AVSD: for instance, in the 10°-AVSD experiment, the pose error of tophat 1 is reduced from 9.35° to 4.88°, and the pose error of tophat 2 is reduced from 21.16° to 13.22°. These results provide further validation of the framework for accommodating obstruction that was introduced in Chapter 6.

7.3 Conceptual Approaches to Dealing with Noncanonicity

We seek an expansion to our framework that will enable accommodation of noncanonical responses from modeled primitives. In theory, we could accommodate almost any type of noncanonical response from a modeled primitive by expanding the primitive parameterization, modifying the measurement model, or applying some combination

component	P_b	10° AVSD			20° AVSD		
		$\hat{\theta}_i^{\mathbf{p}}$ az/el rmse [°]	$\hat{\theta}_i^a$ error [dBsm] mean	stdev	$\hat{\theta}_i^{\mathbf{p}}$ az/el rmse [°]	$\hat{\theta}_i^a$ error [dBsm] mean	stdev
tophat 1	0.00	9.35	16.17	0.40	7.34	16.40	0.47
	0.25	4.88	16.65	0.32	4.73	16.54	0.47
tophat 2	0.00	21.16	18.05	2.02	25.40	18.34	1.91
	0.25	13.22	18.98	0.80	13.48	19.08	1.10

Table 7.4. Comparison of tophat pose and base amplitude estimate statistics with and without obstruction-modeling capability.

of these two approaches. In particular, if we wished to enable explicit description of the variation in a primitive response attributable to a certain kind of noncanonicity, we could expand $\boldsymbol{\theta}_i$ to include relevant descriptors related to that variation. Alternatively, if we wished to accommodate variations in primitive response without explicitly including additional descriptors in our primitive parameterization, we could modify the measurement model $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$ to express the dependences of $\boldsymbol{\lambda}$ and $\mathbf{Z}$ on certain types of noncanonicity. Both of these approaches have important practical difficulties associated with them. We first examine the accommodation of noncanonicity by expansion of primitive parameterization, and then examine accommodation of noncanonicity by modification of the measurement model.

Using the notation of (3.1), the general form of the parameterization of any primitive is

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_i^t \\ \theta_i^{\mathbf{x}} \\ \boldsymbol{\theta}_i^{\mathbf{d}} \end{bmatrix}, \tag{7.6}$$

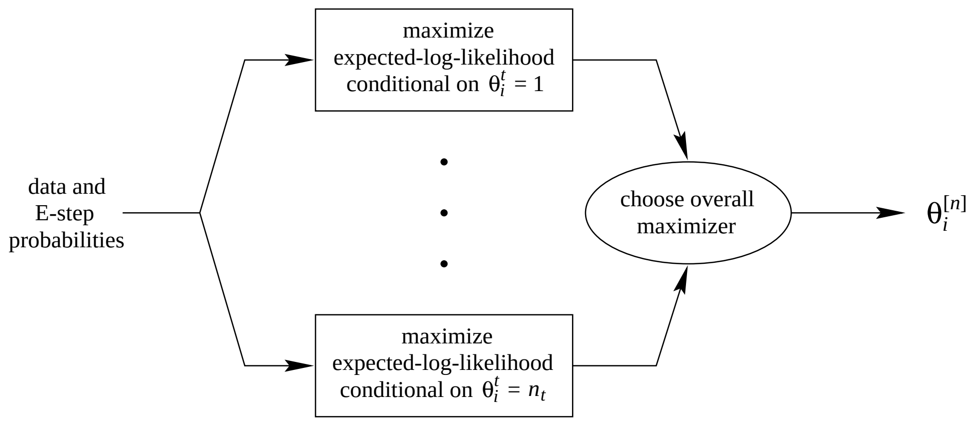
where $\boldsymbol{\theta}_i^{\mathbf{d}}$ is a vector of primitive descriptors, possibly unique to each primitive type. Expansion of the primitive parameterization to model certain noncanonicities explicitly simply corresponds to augmentation or modification of $\boldsymbol{\theta}_i^{\mathbf{d}}$ to include components other than pose, base amplitude, and radius-of-curvature. In order to accommodate an augmented or modified $\boldsymbol{\theta}_i^{\mathbf{d}}$ we would need to modify the existing algorithm in two ways: first, we would need to construct new scattering-response functions $S_{\theta_i^t}(\cdot)$ relating the augmented primitive parameter vector to observed primitive responses, and second, we would need to modify the M step to enable estimation of $\boldsymbol{\theta}_i^{\mathbf{d}}$. The first modification would likely have only a cosmetic effect on the implementation of the algorithm, as long as $S_{\theta_i^t}(\cdot)$ could still be expressed in a convenient functional form or in terms of a previously specified library of basis functions. The second modification, on the other hand, could greatly complicate the implementation of the M step. In particular, the M step requires maximization of each $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\boldsymbol{\theta}_i$. This is accomplished as n_t candidate maximizations (one for each possible θ_i^t) of $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$. Even with $\boldsymbol{\theta}_i^{\mathbf{d}}$ comprising only two parameters (pose and base amplitude) this maximization represents the computational bottleneck of our existing algorithm (see Section 4.7.1). Augmentation of $\boldsymbol{\theta}_i^{\mathbf{d}}$ to include other descriptors—even for a single primitive type—would greatly increase the computational burden of each M step unless the implementation of this stage of the algorithm was fundamentally modified.

A second approach to accommodating noncanonicity is to modify the measurement model $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$ so that variations in $\boldsymbol{\lambda}$ and $\mathbf{Z}$ associated with particular types of noncanonicities are explicitly described, but without reliance on additional primitive descriptors. This approach avoids complicating the M step with explicit estimation of additional primitive descriptors, but still conceptually enables the correction of the gross errors in pose estimates induced by noncanonical responses. The practical difficulty with this approach is the problem of how to construct such a measurement model. For instance, in theory $p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})$ could be modified to capture correlations between amplitude errors at different viewing angles associated with certain types of noncanonicity. In practice, however, learning a pdf to describe a complex set of dependences between elements of a high-dimensional collection of data is often problematic [30, 107]. In order to learn an accurate new measurement model we would require the availability of a large library of characteristic responses or a detailed analytical understanding of how primitive responses vary with the type of noncanonicity under consideration. In either case, we might also need to impose significant regularizing assumptions in order to learn such a pdf [30, 107].

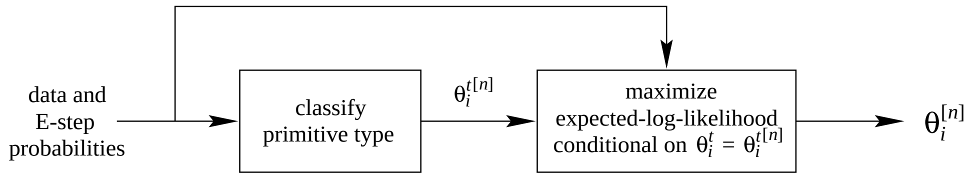
Thus both of these approaches have practical difficulties associated with them. Our approach to accommodating noncanonical primitive responses is similar to the first approach described here, but it relies on a basic change to the implementation of the M step to limit the effect of augmentation of $\boldsymbol{\theta}_i^{\mathbf{d}}$ on the conceptual burden of the algorithm. We now describe this approach.

■ 7.4 Our Approach

We accommodate noncanonical primitive responses by introducing additional primitive descriptors that capture and describe the variations in primitive response associated with the type of noncanonicity under consideration. This primitive-descriptor augmentation is performed in conjunction with a fundamental modification to the M step that limits the computational effect on the algorithm. In particular, the modified M step separates the process of type estimation from that of the estimation of all other components of $\boldsymbol{\theta}_i$. This allows the M step to be implemented for each primitive not as a series of n_t candidate maximizations over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$ for each primitive type, but instead as a two-stage process requiring only a single maximization over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$ for a fixed primitive type. In the first stage of this modified M step a type estimate is obtained from $\mathbf{Z}$ and the set of $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$, without consideration of $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$. In the second stage, $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ is maximized over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$ conditional on the type estimate just obtained. This has the important effect of limiting the increase in computational complexity associated with the augmentation of any single primitive



old implementation



new implementation

Figure 7.7. Block diagrams of the implementation of the M step. The top block diagram illustrates the original implementation of the M step; the bottom block diagram illustrates the proposed new modified implementation of the M step. Although the original implementation requires n_t separate maximizations of the expected-log-likelihood function for each primitive, the proposed modified implementation requires only one.

type's $\boldsymbol{\theta}_i^{\mathbf{d}}$. Figure 7.7 depicts two block diagrams that illustrate the difference between the original implementation and the proposed new implementation of the M step. (The details of the type classification forming the first block in the modified implementation are presented shortly.)

The approach we propose is general enough to accommodate a wide variety of noncanonicities: conceptually, no restriction is placed on $\boldsymbol{\theta}_i^{\mathbf{d}}$, although implementation of the second (type-conditional maximization) stage of our modified M step will depend on the primitive descriptors chosen. As indicated previously, we concentrate on the implementation of this approach for the accommodation of the specific type of geometrical noncanonicity observed from SLICY's tophats and trihedrals. We describe our augmented $\boldsymbol{\theta}_i^{\mathbf{d}}$ and the related type-conditional maximization in Sections 7.4.4 and 7.4.5, respectively. We first describe how the prerequisite type classification is achieved.

7.4.1 Type Classification

Conceptually, the optimal way to estimate the collection of primitive types

$$\boldsymbol{\theta}^t = \begin{bmatrix} \theta_1^t & \cdots & \theta_N^t \end{bmatrix} \tag{7.7}$$

from $\mathbf{Z}$ is to take $\boldsymbol{\theta}^t$ to be the maximum likelihood estimator, *i.e.*, to find

$$\hat{\boldsymbol{\theta}}^t = \arg\max_{\boldsymbol{\theta}^t} p(\mathbf{Z}|\boldsymbol{\theta}^t). \tag{7.8}$$

This approach is intractable for a variety of reasons, chiefly the high dimension of $\mathbf{Z}$ and the mixture-density form of $p(\mathbf{Z}|\boldsymbol{\theta})$ that motivated the introduction of $\boldsymbol{\lambda}$ into the measurement model in the first place. A more practical approach to obtaining $\boldsymbol{\theta}^t$ would be to take

$$\hat{\boldsymbol{\theta}}^t = \arg\max_{\boldsymbol{\theta}^t} p(\hat{\boldsymbol{\lambda}}, \mathbf{Z}|\boldsymbol{\theta}^t) \tag{7.9}$$

for some suitable estimate $\hat{\boldsymbol{\lambda}}$ of $\boldsymbol{\lambda}$. For instance, we could set

$$\hat{\lambda}_{k,j} = \begin{cases} \arg\max_i \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}), & \max_i \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) > \frac{1}{2}, \\ 0, & \max_i \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \leq \frac{1}{2}. \end{cases} \tag{7.10}$$

(Such a definition would ensure that $\hat{\boldsymbol{\lambda}}$ is a feasible vector of label parameters, *i.e.*, that each $\boldsymbol{\lambda}_k$ contains no duplicated nonzero elements.) The independence assumptions of Section 3.3 would then allow decoupling of (7.9) into independent type classifications for each of the N primitives:

$$\hat{\theta}_i^t = \arg\max_{\theta_i^t} p(\hat{\boldsymbol{\lambda}}, \mathbf{Z}|\theta_i^t). \tag{7.11}$$

Maximization of (7.11), however, still requires the availability of the unspecified pdf $p(\hat{\boldsymbol{\lambda}}, \mathbf{Z}|\theta_i^t)$. This pdf could be difficult to obtain from the existing measurement model due to the high dimension of $\mathbf{Z}$ and $\hat{\boldsymbol{\lambda}}$ and the complicated nature of the dependence of these quantities on the non-type elements of $\boldsymbol{\theta}_i$.

Our approach to type classification is to obtain $\hat{\theta}_i^t$ not by maximizing the likelihood $p(\hat{\boldsymbol{\lambda}}, \mathbf{Z}|\theta_i^t)$ as in (7.11), but instead by maximizing a much lower-dimensional likelihood function corresponding to the pdf of an appropriately chosen *summary statistic* $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$. In particular, we take

$$\hat{\theta}_i^t = \arg\max_{\theta_i^t} p(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t). \tag{7.12}$$

For a low-dimensional summary statistic $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$, the pdf $p(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t)$ will be much easier to specify functionally or to learn empirically than $p(\hat{\boldsymbol{\lambda}}, \mathbf{Z}|\theta_i^t)$. The conceptual challenge in performing type classification as in (7.12) is the selection of an appropriate summary statistic that provides a low-dimensional but still informative summary of $\hat{\boldsymbol{\lambda}}$ and $\mathbf{Z}$ that can be used to distinguish between primitive types.

7.4.2 Selection of a Summary Statistic

The ideal summary statistic would be a low-dimensional representation of $\mathbf{Z}$ and $\hat{\boldsymbol{\lambda}}$ that takes on markedly different values for primitives of different types, but which is nearly invariant to differences in primitive pose, location, base amplitude, and noncanonicity *within* any particular type. Several features of $\mathbf{Z}$ and $\hat{\boldsymbol{\lambda}}$ meet both of these criteria. For instance, the set of binary polarimetric-signature classifications associated with each primitive, $\{t_{k,j} : \hat{\lambda}_{k,j} = i\}$, is not only informative in identifying the type of primitive i, but is also relatively invariant to variations within a type class. Similarly informative is the pattern of detections and missed detections on the viewsphere as provided by $\hat{\boldsymbol{\lambda}}$. In particular, each primitive type produces reports over a characteristic region of the viewsphere: tophats produce a large number of reports spread across the entire viewsphere, trihedrals produce reports spread across a single octant of the viewsphere, and dihedrals and cylinders produce a small number of reports near a single great circle on the viewsphere. Furthermore, these patterns are not dramatically affected by variations within a type class: they are essentially invariant to changes in primitive location and radius; they display limited but not fundamental variation with changes in base amplitude; they vary in orientation but not in shape and extent with primitive pose; and they are similar for all but the most severe types of noncanonicity.

In comparison to the these sources of information, the other features of $\hat{\boldsymbol{\lambda}}$ and $\mathbf{Z}$ are less useful in discriminating between primitive type. The set of report locations associated with any primitive, $\{\mathbf{x}_{k,j} : \hat{\lambda}_{k,j} = i\}$, provides little distinguishing information, and although the set of report amplitudes $\{a_{k,j} : \hat{\lambda}_{k,j} = i\}$ is more informative, it is extremely sensitive to variations in base amplitude (obviously) and to noncanonicity. We thus construct our summary statistic based on $\{t_{k,j} : \hat{\lambda}_{k,j} = i\}$ and on the pattern of detections provided by $\hat{\boldsymbol{\lambda}}$. In particular, our $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ comprises four scalar elements, the first of which summarizes the information provided by $\{t_{k,j} : \hat{\lambda}_{k,j} = i\}$ and the remaining three of which describe $\hat{\boldsymbol{\lambda}}$.

The first element of our $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ summarizes $\{t_{k,j} : \hat{\lambda}_{k,j} = i\}$ by describing the fraction of these binary measurements which correspond to odd-bounce polarimetric-signature classifications:

$$f_t = \frac{\#\{t_{k,j} : \hat{\lambda}_{k,j} = i,\, t_{k,j} = 1\}}{\#\{t_{k,j} : \hat{\lambda}_{k,j} = i\}}, \tag{7.13}$$

where # indicates the cardinality of the set it is applied to. (High f_t values tend to indicate trihedrals or cylinders, while low f_t values tend to indicate tophats or dihedrals.)

The second element of $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ is simply the fraction of images in the data set in which the primitive generated a report:

$$f_d = \frac{\#\{\hat{\lambda}_{k,j} : \hat{\lambda}_{k,j} = i\}}{K}. \tag{7.14}$$

The remaining two components of $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ describe the spread of detections on the viewsphere in terms of the eigenvalues of the *orientation matrix* [2, 34, 108, 114] associated with their viewing angles. In particular, suppose that $\hat{\boldsymbol{\lambda}}$ indicates N_d detections of primitive i, at viewing elevations and azimuths given by $\{(\psi_n, \phi_n)\}_{n=1}^{N_d}$. Each of these viewing angles can be specified in terms of a unit vector in $\mathbb{R}^3$ extending from the origin to a point on the unit sphere:

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = \begin{bmatrix} \cos\psi_n \cos\phi_n \\ -\cos\psi_n \sin\phi_n \\ \sin\psi_n \end{bmatrix}. \tag{7.15}$$

The orientation matrix is defined in terms of these unit vectors as

$$\mathbf{M} = \begin{bmatrix} \sum x_n^2 & \sum x_n y_n & \sum x_n z_n \\ \sum x_n y_n & \sum y_n^2 & \sum y_n z_n \\ \sum x_n z_n & \sum y_n z_n & \sum z_n^2 \end{bmatrix}, \tag{7.16}$$

where each summation is over $n = 1, \ldots, N_d$. The eigenvalues and eigenvectors of $\mathbf{M}$ provide a concise and informative description of the distribution and orientation of directional data on the sphere [2,34,108,114].[4] In particular, the relative magnitudes of the eigenvalues of $\mathbf{M}$ are often used to summarize the distribution of directional data on the sphere [34, 114, 115]: similar magnitudes for all eigenvalues indicate data spread nearly uniformly across the viewsphere, while different ratios between the eigenvalues are characteristic of data concentrated near one direction or along a great circle.

Let us denote the eigenvalues of $\mathbf{M}$ by v_1, v_2, and v_3, where $v_1 \geq v_2 \geq v_3$. Figure 7.8 presents scatter plots of the variation in $\log_{10}(v_2/v_3)$ *vs.* $\log_{10}(v_1/v_3)$ for each primitive type. Each point in each plot was calculated by taking a randomly selected 20°-AVSD data set corresponding to a randomly oriented 10-dBsm primitive of the specified type, and simulating the detection of that primitive using the canonical scattering-response function $S_{\theta_i^t}(\cdot)$ and the probability-of-detection function $P_D(\cdot)$ as described in Section 5.2. Figure 7.8 indicates that although there is some variation in these eigenvalue ratios within each primitive type, these statistics are informative in discriminating between primitive types. We thus use scaled versions of $\log_{10}(v_2/v_3)$ and $\log_{10}(v_1/v_3)$

[4] If $\{(\psi_n, \phi_n)\}_{n=1}^{N_d}$ are envisioned as the coordinates of a set of N_d unit point masses on the unit sphere, the eigenvectors of $\mathbf{M}$ have the following physical interpretation: the eigenvector corresponding to the maximum eigenvalue represents the axis about which the moment of inertia is minimized, and the eigenvector corresponding to the minimum eigenvalue represents the axis about which the moment of inertia is maximized [108, 114]. The remaining eigenvector provides an orthogonal completion of $\mathbb{R}^3$.

as the final two elements of our summary statistic:

$$\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z}) = \begin{bmatrix} f_t \\ f_d \\ \frac{1}{8}\log_{10}\frac{v_1}{v_3} \\ \frac{1}{8}\log_{10}\frac{v_2}{v_3} \end{bmatrix}. \tag{7.17}$$

(The 1/8 scaling coefficients we apply to the log-ratios of the orientation-matrix eigenvalues ensure that each component of $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ varies over a similar range, *i.e.*, approximately from 0 to 1. The motivation for this will be explained shortly.)

7.4.3 Implementation of Summary-Statistic Type Classification

Classification of type as in (7.12) requires the availability of the pdf $p(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t)$. We obtain this pdf using a Parzen-window density estimator (PWDE) [84,107]. The PWDE provides an estimate of a given pdf in terms of a set of samples from that pdf. In particular, given a set of samples $\{\mathbf{Y}_n\}_{n=1}^{N_s}$ obtained from a pdf $p(\mathbf{y})$, the PWDE of $p(\mathbf{y})$ is of the form[5]

$$\hat{p}(\mathbf{y}) = \frac{1}{N_s}\sum_{n=1}^{N_s} \kappa\left(\frac{1}{h}\|\mathbf{y} - \mathbf{Y}_n\|\right) \cdot \frac{c_\kappa}{h^d}, \tag{7.18}$$

where $\kappa(.)$ is some appropriately chosen kernel function, h is an appropriately chosen "bandwidth parameter" determining the spread of the kernel function around each sample value, d is the dimension of $\mathbf{y}$, and c_κ is a normalization constant ensuring

$$\frac{c_\kappa}{h^d}\int \kappa\left(\frac{\mathbf{y}}{h}\right) d\mathbf{y} = 1. \tag{7.19}$$

There is a considerable body of literature describing appropriate selection of kernel functions and bandwidth parameters [30, 40, 83, 107]. In situations where each element of the sample vector varies over a similar range and there are few isolated sample points, a common choice for $\kappa(.)$ is a circularly symmetric Gaussian kernel. Because the elements of our summary statistic all vary roughly between 0 and 1,[6] and because Figure 7.8 suggests that samples of $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ will contain few outliers, we employ such a kernel to obtain a PWDE of the form

$$\hat{p}(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t) = \frac{1}{N_s}\sum_{n=1}^{N_s} \frac{1}{(2\pi)^2 h^4}\exp\left(-\frac{1}{2h^2}\|\mathbf{s} - \mathbf{S}_n(\theta_i^t)\|\right), \tag{7.20}$$

[5] Other, more general forms of the Parzen-window density estimator also exist. For instance, in one common implementation, the bandwidth parameter h varies at each sample point depending on the proximity to other samples.

[6] This is the motivation for scaling $\log_{10}(v_2/v_3)$ and $\log_{10}(v_1/v_3)$ by 1/8 in the definition of $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ in (7.17).

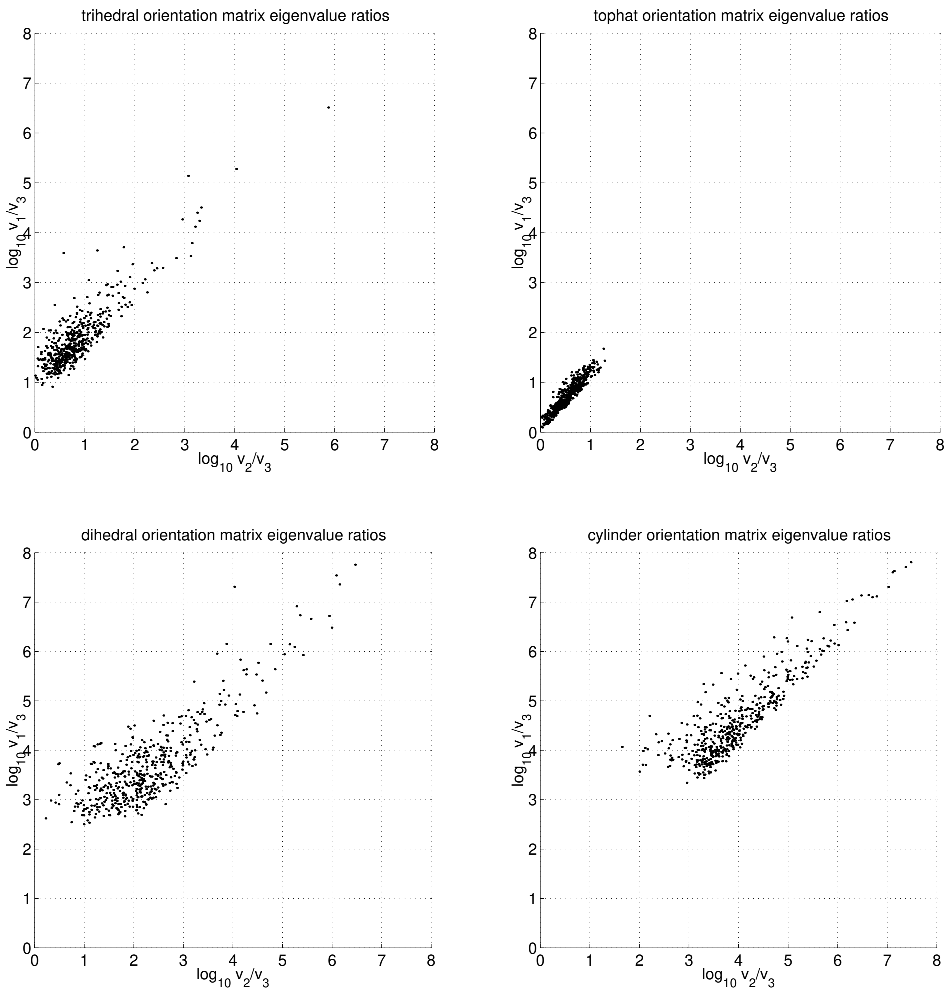


Figure 7.8. Orientation matrix eigenvalue ratios. These plots characterize the variability of the spread of detections on the viewsphere in a 20°-AVSD data set for randomly oriented 10-dBsm trihedrals (upper left), tophats (upper right), dihedrals (lower left), and cylinders (lower right). The vertical axis of each plot represents the log-ratio between the maximum and minimum eigenvalues of the orientation matrix; the horizontal axis represents the log-ratio between the middle and minimum eigenvalues of the matrix.

where $\{\mathbf{S}_n(\theta_i^t)\}_{n=1}^{N_s}$ is a set of samples of $p(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t)$ obtained for a fixed value of θ_i^t. We generate these samples by simulating the detection process for each primitive type using $S_{\theta_i^t}(\cdot)$ and $P_D(\cdot)$ as described in Section 7.4.2, and by simulating the generation of $t_{k,j}$ measurements for each detection according to the confusion probabilities provided by the confusion matrix $\boldsymbol{\rho}$ of Table 5.2. We choose h using a standard criterion in classification-related PWDE applications [30, 35]: we construct an empirical pdf as in (7.20) for many different values of h, and choose h to be the value that achieves the best classification performance on an additional set of samples randomly generated in a manner identical to the first set. We choose $h = 0.0375$ based on such an analysis.[7]

Given an empirical pdf obtained as just described, the first component of our modified M step is extremely simple: $\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$ is calculated from $\mathbf{Z}$ and the set of $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ for each primitive. The primitive type estimate $\theta_i^{t[n]}$ is then taken as

$$\theta_i^{t[n]} = \arg\max_{\theta_i^t} \hat{p}(\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})|\theta_i^t). \tag{7.21}$$

The second component of the M step then entails maximization of $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$, conditional on the type classification obtained in (7.21). We now describe our augmented $\boldsymbol{\theta}_i^{\mathbf{d}}$ and the related maximization.

7.4.4 Selection of an Augmented Primitive-Descriptor Vector

Any choice of $\boldsymbol{\theta}_i^{\mathbf{d}}$ should meet two criteria: first, it should enable accurate description of primitive responses over the range of noncanonicities we wish to model (*i.e.*, variations in trihedral and tophat responses of the form observed in Section 7.2); second, it should consist of descriptors that can be reliably and robustly estimated from $\mathbf{Z}$, otherwise the augmented parameterization will offer no additional benefit to a model-based ATR system. For this reason we select a relatively coarse-grained augmentation to the existing $\boldsymbol{\theta}_i^{\mathbf{d}}$ that takes the form of a discrete variable θ_i^s characterizing each primitive as one of a number of discrete *subtypes*, *i.e.*, variants on the basic canonical primitive type that exhibit different responses. In particular, each $\boldsymbol{\theta}_i$ is now of the form

$$\boldsymbol{\theta}_i = \begin{bmatrix} \theta_i^t \\ \theta_i^{\mathbf{x}} \\ \theta_i^s \\ \theta_i^{\mathbf{P}} \\ \theta_i^a \\ \theta_i^r \end{bmatrix}, \tag{7.22}$$

where θ_i^s is a discrete subtype-indicator index for a primitive of type θ_i^t that take on values $1, \ldots, n_s(\theta_i^t)$. Although finer-grained parameterizations are possible, we select

[7] This value of h attained a misclassification probability of 0.0055. Other values tested in the range $0.0025 \leq h \leq 0.125$ gave misclassification probabilities not exceeding 0.017, suggesting that type-classification performance is relatively robust to the choice of h.

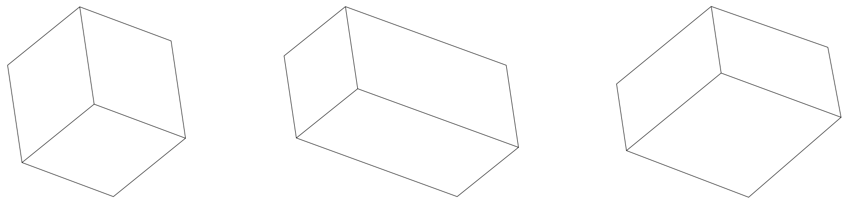


Figure 7.9. Trihedral subtype geometries. Subtype 1 is the canonical trihedral; subtype 2 represents a trihedral extended along one axis; subtype 3 represents a trihedral with an extended base.

this approach because it provides a simple illustration of our framework and because it provides a benchmark for the selection of more detailed choices of $\boldsymbol{\theta}_i^{\mathbf{d}}$: if θ_i^s cannot be estimated reliably, it is unlikely that finer-grained parameterizations could be estimated reliably without inclusion of additional features in $\mathbf{Z}$. In Chapter 8 we discuss approaches to the selection of $\boldsymbol{\theta}_i^{\mathbf{d}}$ to model observed variations in primitive responses.

We utilize a set of three trihedral subtypes and four tophat subtypes that are representative of the types of noncanonical responses observed from SLICY. In particular, the trihedral subtypes are intended to represent variations in the relative size of the trihedral plate components, as depicted in Figure 7.9: the first subtype is the canonical square-plate trihedral considered in previous chapters, the second subtype corresponds to a trihedral that has been extended along one of its axes, and the third subtype response corresponds to a trihedral with an extended base, *i.e.*, a trihedral that has been extended along two of its axes. Each of these trihedrals can be uniquely oriented by specification of three Euler angles in $\theta_i^{\mathbf{p}}$, as depicted in Figure 7.10. We employ a set of heuristic scattering-response functions to model the angular scattering responses from these types of trihedrals. The construction and functional forms of these subtype responses are detailed in Appendix C; the three responses themselves are depicted in Figure 7.11. (The poses and base amplitudes of each of the three responses in Figure 7.11 are identical.) In the context of SLICY, the hope is that the second trihedral subtype response might enable more accurate modeling of trihedrals 1 and 2, while the third trihedral subtype response might enable more accurate modeling of trihedral 3.

We utilize a set of four tophat subtypes that correspond to variations in the size and shape of the tophat base, as depicted in Figure 7.12. In particular, the first tophat subtype is identical to the canonical tophat response of earlier chapters. The second and third subtype responses are intended to model the responses from tophats with different-sized square bases; the final subtype response is meant to model the response from a tophat asymmetrically centered along one axis of an extended rectangular base. Each of these four tophats can be uniquely oriented by specification of three Euler angles in $\theta_i^{\mathbf{p}}$, as depicted in Figure 7.13.[8] The scattering response we use for each subtype is

[8]Because the first tophat subtype corresponds to the rotationally symmetric canonical tophat geometry, only two Euler angles are needed to define its orientation, as described in Section 3.1.

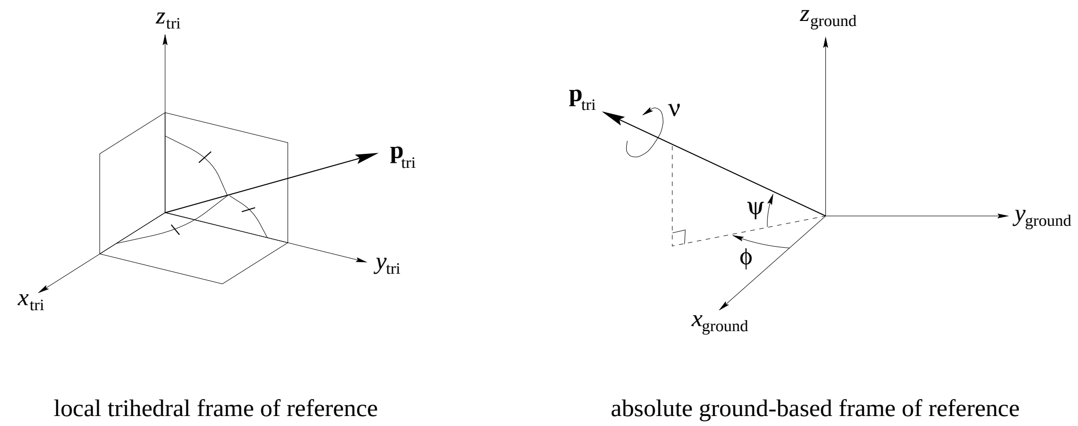


Figure 7.10. Specification of the orientation of trihedral subtypes. The three axes of a trihedral define a local frame of reference, as depicted on the left; a trihedral "pose vector" $\mathbf{p}_{\mathrm{tri}}$ can be defined in this local frame of reference as the direction equidistant from the three trihedral axes (*i.e.*, the direction separated from each axis by 54.74° on the viewsphere). Trihedral pose in the absolute ground-based frame of reference can then be defined in terms of three Euler angles, as depicted on the right: an elevation ψ and azimuth ϕ corresponding to the direction of $\mathbf{p}_{\mathrm{tri}}$ in terms of the ground-based axes, and a rotation angle ν representing the rotation of the trihedral axes about $\mathbf{p}_{\mathrm{tri}}$. (We define $\nu = 0°$ to correspond to the rotation at which the dot product between the local and absolute z axes is maximized.)

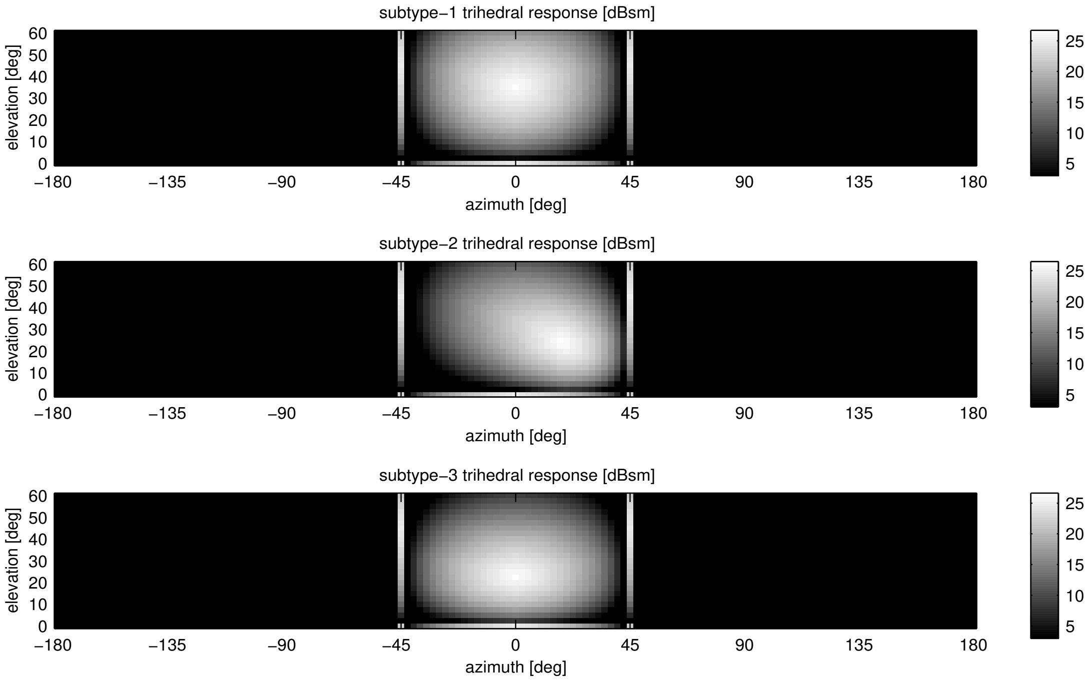


Figure 7.11. Trihedral subtype responses. These are the models used for the three trihedral subtypes depicted in Figure 7.9.

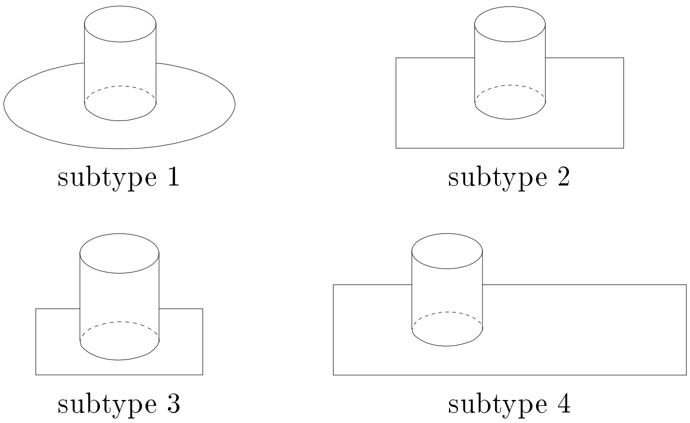


Figure 7.12. Tophat subtype geometries. Subtype 1 is the canonical tophat; subtype 2 represents a square-base tophat with base extent in cardinal directions equal to the tophat height; subtype 3 represents a square-base tophat with smaller base; subtype 4 represents a tophat with a rectangular base extended in one direction.

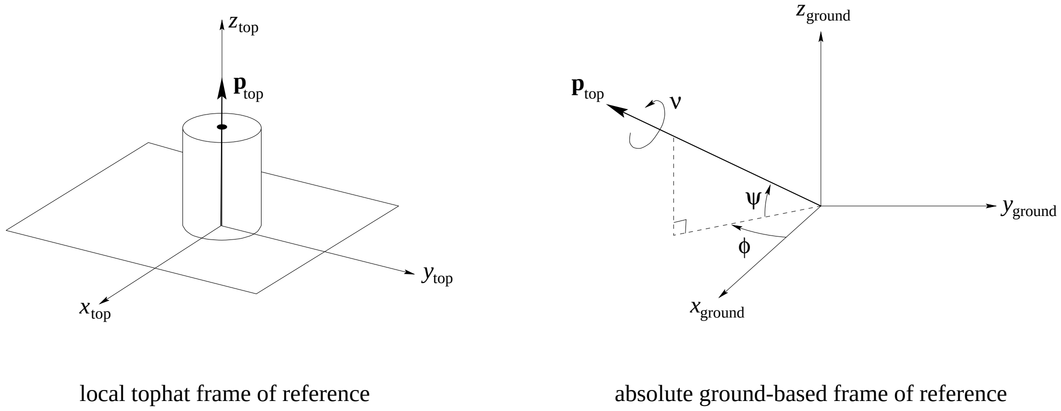


Figure 7.13. Specification of the orientation of tophat subtypes. The cylinder axis and cardinal base directions of a tophat define a local frame of reference, as depicted on the left; a tophat "pose vector" $\mathbf{p}_{\mathrm{top}}$ can be defined in this local frame of reference as the cylinder axis direction. Tophat pose in the absolute ground-based frame of reference can then be defined in terms of three Euler angles, as depicted on the right: an elevation ψ and azimuth ϕ corresponding to the direction of $\mathbf{p}_{\mathrm{top}}$ in the ground-based frame of reference, and a rotation angle ν representing the rotation of the tophat axes about $\mathbf{p}_{\mathrm{top}}$. (We define $\nu = 0^\circ$ to correspond to the rotation at which the dot product between the local and absolute x axes is maximized.)

depicted in Figure 7.14. (The poses and base amplitudes of each of the four responses in Figure 7.14 are identical.) As with the trihedral subtype responses, the construction and functional forms of the tophat subtype responses are detailed in Appendix C.

Dihedrals and cylinders are each modeled with a single subtype response that corresponds to the canonical response from each primitive. This is motivated by the low observabilities of these two primitive types.

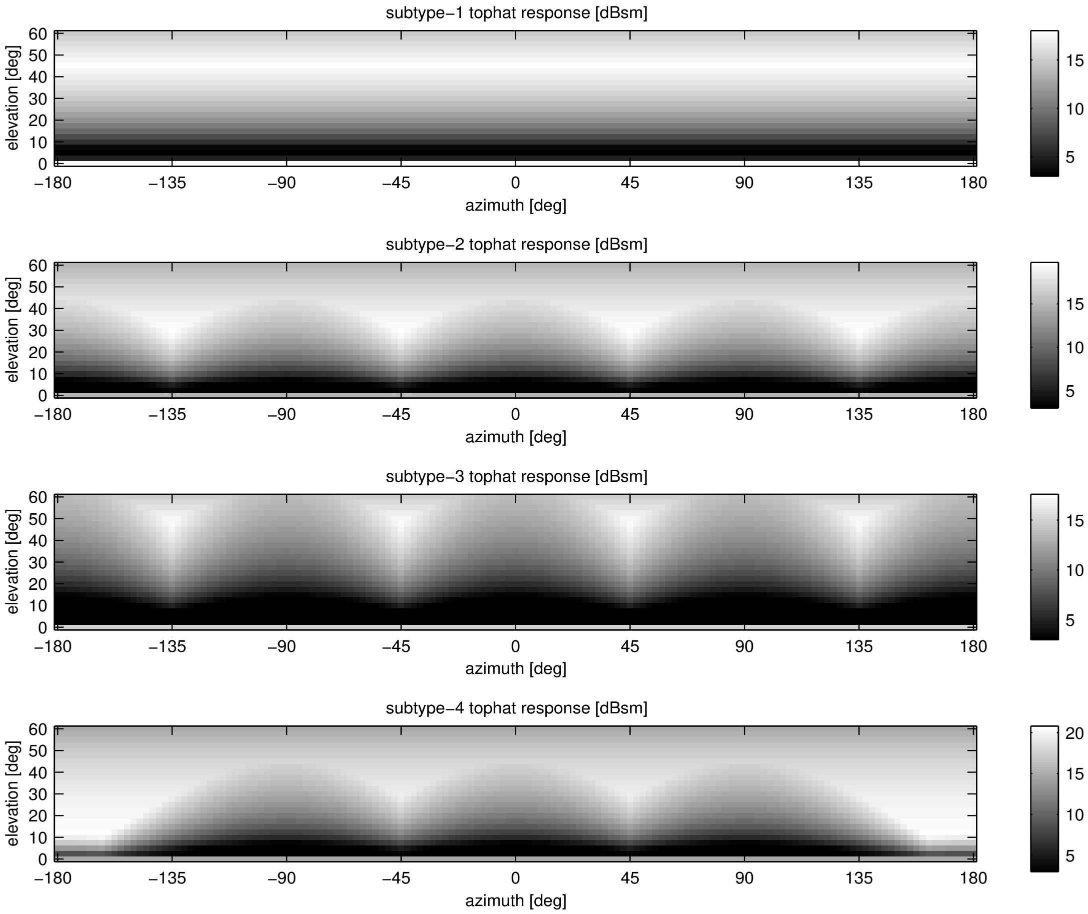


Figure 7.14. Tophat subtype responses. These are the models used for the four tophat subtypes depicted in Figure 7.12.

7.4.5 Implementation of Type-Conditional Maximization

Given the inclusion of subtype in the primitive-descriptor vector, the second component of the modified M step—*i.e.*, the type-conditional maximization of $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\theta_i^{\mathbf{x}}$ and $\boldsymbol{\theta}_i^{\mathbf{d}}$—can be achieved in a manner similar to that of the original M step. In particular, given the type classification $\theta_i^{t[n]}$, we first maximize $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ over $\theta_i^{\mathbf{x}}$ and, if required, θ_i^r in closed form according to (4.28) or (4.29). We then maximize over $\theta_i^s, \theta_i^{\mathbf{P}}$, and θ_i^a by performing $n_s(\theta_i^{t[n]})$ candidate maximizations over pose and base amplitude, one for each subtype. We then compare the candidate maximizers for each subtype to choose the overall maximizer. Because the type-classification procedure described in Section 7.4.1 reduces the overall computational burden of the M step by a factor of roughly n_t, we perform each candidate subtype maximization jointly over $\theta_i^{\mathbf{P}}$ and θ_i^a. (This differs from the maximization procedure of Section 4.3, in which $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ was maximized over θ_i^a and $\theta_i^{\mathbf{P}}$ sequentially according to the ECM method.)

To demonstrate the performance of the modified algorithm on a target consisting of canonical primitives, we performed 200 runs of the modified algorithm on the target depicted in Figure 5.5 at AVSDs of 10° and 20°. The performance of the modified algorithm on this target is presented in Table 7.5, along with the results from the existing algorithm (originally reported in Table 6.9). The statistics for the two algorithms are extremely similar: for every primitive at each AVSD, P_{disc} and P_{id} are almost identical between algorithms, indicating that the modified implementation of the M step is not markedly affecting the ability of the algorithm to discover and correctly identify primitives. The modification of the maximization procedure to estimate $\theta_i^{\mathbf{P}}$ and θ_i^a jointly, rather than sequentially, has a slight effect on the error statistics associated with these parameters. In particular, some of the base amplitude estimate biases change slightly (*e.g.*, in the 20°-AVSD experiment the trihedral base amplitude estimate bias changes from −0.84 dBsm to −0.40 dBsm and the dihedral base estimate bias changes from −2.87 dBsm to −1.56 dBsm), and some of the pose errors change slightly (*e.g.*, in the 10°-AVSD experiment the trihedral pose rotational error changes from 15.20° to 7.23°, and in the 20°-AVSD experiment the dihedral pose rotational error changes from 14.91° to 20.29°). Overall, however, Table 6.9 demonstrates similar performance for the two algorithms, implying little loss in performance of this modified algorithm for problems for which the original algorithm was designed.

In addition to the quantities presented in Table 6.9, the modified algorithm provides subtype estimates. Table 7.6 presents the fraction of runs on this target at each AVSD in which the trihedral and tophat were classified as each subtype. (Recall that a single subtype is used for the dihedral and cylinder.) This table demonstrates that the modified algorithm is correctly classifying the primitives as the canonical subtype in almost all algorithm runs: in the 10°-AVSD experiment, the tophat was classified as canonical in every run and the trihedral was classified as canonical in all but one run. In the 20°-AVSD experiment, the tophat subtype was correctly classified in 93% of the runs, and the trihedral was correctly classified as canonical in 89% of the runs.

			P_{disc}	P_{id}	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
					bias	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tri	old	1.00	1.00	−0.43	0.42	7.60	2.41	1.94	15.20	—
		new	1.00	1.00	−0.28	0.36	7.49	2.50	1.46	7.23	—
	top	old	1.00	1.00	−0.76	0.18	0.51	1.36	1.45	—	1.11
		new	1.00	1.00	−0.80	0.16	0.33	1.39	1.38	—	1.11
	dih	old	0.96	1.00	−0.64	2.09	5.73	12.83	11.40	6.84	—
		new	0.95	1.00	−0.09	1.71	6.44	12.24	9.72	7.22	—
	cyl	old	0.40	0.99	1.14	3.24	2.22	18.97	2.62	—	6.22
		new	0.36	1.00	1.65	4.61	0.94	21.55	3.73	—	7.00
20°	tri	old	1.00	1.00	−0.84	0.93	7.37	4.43	4.05	23.04	—
		new	1.00	1.00	−0.40	0.82	7.07	4.87	4.11	23.95	—
	top	old	1.00	1.00	−0.75	0.30	0.44	2.59	2.37	—	1.94
		new	1.00	1.00	−0.82	0.44	0.17	2.47	2.57	—	1.97
	dih	old	0.49	0.98	−2.87	4.04	6.63	23.75	15.40	14.91	—
		new	0.56	0.96	−1.56	4.15	4.87	21.74	13.23	20.29	—
	cyl	old	0.11	0.94	2.95	5.46	4.30	34.37	4.27	—	10.19
		new	0.15	0.93	1.25	4.12	6.41	37.24	3.96	—	8.74

Table 7.5. Performance of original algorithm and modified algorithm on the target of Section 5.4.1. Here "old" refers to the algorithm developed in previous chapters (and incorporating the modification to the initialization procedure described in Section 7.2.1) and "new" refers to the algorithm incorporating all modifications described in this chapter.

component	subtype classification fraction	
	10° AVSD	20° AVSD
tophat	$[\,1.000 \quad 0.000 \quad 0.000 \quad 0.0000\,]$	$[\,0.930 \quad 0.050 \quad 0.015 \quad 0.005\,]$
trihedral	$[\,0.995 \quad 0.005 \quad 0.000\,]$	$[\,0.890 \quad 0.065 \quad 0.045\,]$

Table 7.6. Subtype-estimation performance of modified algorithm on the target of Section 5.4.1.

■ 7.5 Performance of the Modified Algorithm on SLICY

We performed 150 runs of the modified algorithm on SLICY at AVSDs of 10° and 20°, using the same algorithmic and measurement-model parameters that were used in Section 7.2.2 and in previous chapters. (As in Section 7.2.2, the obstruction-probability prior was set to $P_b = 0.25$.) The modified algorithm discovered SLICY's dominant scatterers (*i.e.*, its tophats and trihedrals) in each run at each AVSD. Table 7.7 presents the number of runs in which each component of SLICY was discovered, along with the results from the original algorithm (originally presented in Table 7.2). The discoverability of each component is nearly identical between algorithms. In addition to the components indicated in Table 7.7, one or more spurious estimates were produced in most algorithm runs. In particular, an average of 2.68 spurious primitive estimates were produced in each 10°-AVSD run, and an average of 2.07 spurious primitive estimates were produced in each 20°-AVSD run. This is slightly more than were generated by the old algorithm (*i.e.*, an average of 2.23 in the the 10°-AVSD runs and an average of 1.90 in the 20°-AVSD runs). This is attributable to the availability of additional scattering models (*i.e.*, the trihedral and tophat subtype responses), which provide a richer set of responses with which to approximate unmodeled response mechanisms. In Section 7.6 we describe a method for distinguishing between spurious primitive estimates and legitimate primitive estimates produced by explicitly modeled response mechanisms.

Table 7.8 presents the subtype-classification statistics for SLICY's trihedrals and tophats at both AVSDs. The first column for each AVSD, denoted P_{nc}, gives the fraction of runs in which each component was classified as noncanonical (*i.e.*, as a subtype other than 1); the second column gives the confusion statistics between subtypes. We see that in most runs at either AVSD, these components are being correctly classified as noncanonical. Table 7.8 indicates that tophat 1 is usually identified as subtype 2, a square-base subtype whose base extent is similar to its cylinder height; tophat 2 is usually identified as subtype 3, a square-base subtype in which the cylinder height exceeds the base extent (see Figure 7.12). This is commensurate with the different heights of tophats 1 and 2. Despite SLICY's rectangular top base, the tophats are rarely classified as subtype 4. This is because the strong responses expected from such a subtype along the extended base axis are not observed from SLICY's tophats due to obstruction.

Trihedrals 1 and 2 are classified as subtype 2 in most algorithm runs (*i.e.*, in 95% of the 10°-AVSD runs and in 75% of the 20°-AVSD runs). This is commensurate with their geometries, which are similar to that of trihedral subtype 2 (see Figure 7.9). Although the geometry of trihedral 3 is similar to that of subtype 3, Table 7.8 indicates that subtype classification of this trihedral is less successful. In particular, although trihedral 3 is identified as noncanonical in 68% of the 10°-AVSD runs and in 71% of the 20°-AVSD runs, it is classified as subtype 3 in only 47% of the 10°-AVSD runs and 16% of the 20°-AVSD runs. This is attributable to the same factors that complicated the estimation of the parameters of trihedral 3 in Section 7.2.2—namely, its unresolvability from tophat 2 and its proximity to the response mechanisms of dihedrals 3 and 4 and

component	alg	P_{disc} 10° AVSD	P_{disc} 20° AVSD
tophat 1	old	1.00	1.00
	new	1.00	1.00
tophat 2	old	1.00	1.00
	new	1.00	1.00
trihedral 1	old	1.00	1.00
	new	1.00	1.00
trihedral 2	old	1.00	1.00
	new	1.00	1.00
trihedral 3	old	1.00	1.00
	new	1.00	1.00
dihedral 1	old	0.13	0.03
	new	0.17	0.03
dihedral 2	old	0.24	0.04
	new	0.25	0.04
dihedral 3	old	0.41	0.05
	new	0.41	0.07
dihedral 4	old	0.02	0.01
	new	0.02	0.01
cylinder	old	0.05	0.02
	new	0.06	0.02

Table 7.7. Fraction of runs in which each component of SLICY was discovered.

component	10° AVSD P_{nc}	10° AVSD subtype confusion	20° AVSD P_{nc}	20° AVSD subtype confusion
tophat 1	0.86	$[0.14\ 0.83\ 0.03\ 0.00]$	0.73	$[0.27\ 0.59\ 0.08\ 0.06]$
tophat 2	0.98	$[0.02\ 0.00\ 0.97\ 0.01]$	0.95	$[0.05\ 0.03\ 0.90\ 0.02]$
trihedral 1	0.96	$[0.04\ 0.95\ 0.01]$	0.79	$[0.21\ 0.74\ 0.05]$
trihedral 2	0.95	$[0.05\ 0.95\ 0.00]$	0.77	$[0.23\ 0.76\ 0.01]$
trihedral 3	0.68	$[0.32\ 0.21\ 0.47]$	0.71	$[0.29\ 0.55\ 0.16]$

Table 7.8. Subtype-estimation performance of modified algorithm on SLICY.

AVSD	component	alg	$\hat{\theta}_i^a$ error [dBsm]		$\hat{\theta}_i^{\mathbf{x}}$ error [cm]		$\hat{\theta}_i^{\mathbf{P}}$ rmse [°]		$\hat{\theta}_i^r$ rmse
			mean	stdev	‖bias‖	‖stdev‖	az/el	rot	[cm]
10°	tophat 1	old	16.65	0.32	9.57	2.70	4.88	—	3.91
		new	15.49	0.90	8.94	3.44	2.68	3.91	3.80
	tophat 2	old	18.98	0.80	2.35	1.34	13.22	—	0.92
		new	21.49	0.88	1.51	1.06	5.28	3.27	0.63
	trihedral 1	old	27.01	0.80	2.44	1.64	10.44	30.05	—
		new	28.26	0.72	2.43	1.62	4.36	8.97	—
	trihedral 2	old	26.53	0.66	5.83	3.07	8.99	23.33	—
		new	27.57	0.54	7.26	3.28	2.46	6.53	—
	trihedral 3	old	25.07	1.57	8.28	3.15	13.17	26.50	—
		new	26.47	3.93	7.51	3.42	10.18	18.76	—
20°	tophat 1	old	16.54	0.47	10.65	5.56	4.73	—	4.12
		new	15.85	1.51	10.57	5.32	3.18	5.22	4.70
	tophat 2	old	19.08	1.10	2.30	2.61	13.48	—	1.56
		new	21.25	1.48	1.59	1.93	6.80	4.50	1.16
	trihedral 1	old	26.40	1.43	1.83	3.17	10.85	36.71	—
		new	27.86	1.45	1.86	3.20	7.15	23.38	—
	trihedral 2	old	25.97	1.11	5.74	6.24	10.05	29.14	—
		new	27.51	1.07	7.42	6.61	6.25	20.41	—
	trihedral 3	old	25.68	4.69	6.87	8.13	21.95	29.13	—
		new	30.40	6.74	5.74	8.42	20.52	35.20	—

Table 7.9. Results of modified algorithm on SLICY. Here "old" refers to the algorithm developed in previous chapters and "new" refers to the algorithm incorporating the modifications described in this chapter.

the multiple-primitive responses generated by trihedral 3 and tophat 2.

Table 7.9 presents error statistics obtained by the new algorithm on SLICY's tophat and trihedral components. For convenience of comparison, Table 7.9 also lists the error statistics for the old algorithm (*i.e.*, the results presented in Table 7.3). We see that most of the results are similar between algorithms, with several notable exceptions. In particular, the introduction of multiple subtypes has the desired effect of reducing the pose errors associated with most primitives. For instance, the azimuth/elevation pose error of tophat 1 has been reduced from 4.88° to 2.68° in the 10°-AVSD experiment, and from 4.73° to 3.18° in the 20°-AVSD experiment; the reduction for tophat 2 is more dramatic, from 13.22° to 5.28° in the 10°-AVSD experiment and from 13.48° to 6.80° in the 20°-AVSD experiment. Note also that because the three noncanonical tophat subtypes are rotationally asymmetric, their poses include a third Euler angle; Table 7.9 lists the rmse error in this rotational component, conditional on classification as subtype 2, 3, or 4. This error is 5.22° or less for both tophats at both AVSDs, indicating that the algorithm is quite successful at determining the rotational orientation of noncanonical tophats.

The pose errors for trihedral 1 and 2 are also significantly reduced by the modified algorithm. For instance, in the 10°-AVSD experiment, the azimuth/elevation pose error for trihedral 1 has been reduced from 10.44° to 4.36°, and the error for trihedral 2 has been reduced from 8.99° to 2.46°. Additionally, the rotational errors for these primitives have been significantly reduced, due to marked rotational asymmetry of the response of trihedral subtype 2 (see Figure 7.11). For instance, in the 10°-AVSD experiment, the rotational pose error for trihedral 1 has been reduced from 30.05° to 8.97°; the corresponding error for trihedral 2 is reduced from 23.33° to 6.53°.

The modified algorithm has a much less pronounced effect on the pose estimates of trihedral 3: its pose error is reduced from 13.17° to 10.31° in the 10°-AVSD experiment and from 21.95° to 20.92° in the 20°-AVSD experiment. As discussed in Section 7.2.2, the poorer algorithm performance on this primitive is due to the unresolvability of its responses from those of tophat 2 over a range of viewing angles, and its proximity to several other response mechanisms that produce reports at different viewing angles (namely, dihedrals 3 and 4 and the multiple-reflection mechanism generated by trihedral 3 and tophat 2). Because our modified algorithm includes no means for detecting or accommodating such effects, trihedral 3 displays poorer statistics than the other two trihedrals regardless of which algorithm is used. In Chapter 8 we examine how these effects might be accommodated.

Scatter plots of the pose estimates produced for each of the tophats are displayed in Figure 7.15; similar plots for the trihedrals are displayed in Figure 7.16. (As with Figures 7.5 and 7.6, the range of azimuths and elevations displayed in the tophat plots is different from that in the trihedral plots in order to display the variation in pose estimates over the different portions of the viewsphere near each primitive's true pose.) These plots also identify the subtype associated with each pose estimate. Comparison to Figures 7.5 and 7.6 demonstrates that the algorithm is in fact removing much of the gross pose error introduced by noncanonical primitive responses. (Again, trihedral 3 is an exception.) Although residual errors remain, most pose estimates lie closer to their true values. Figures 7.15 and 7.16 also demonstrate that pose errors associated with different subtypes tend to cluster around different values. For instance, the scatter plots for trihedrals 1 and 2 clearly demonstrate a correlation between subtype and pose estimate: although a correct subtype classification for these trihedrals is usually accompanied by a pose estimate near truth, an incorrect subtype classification is likely to be associated with a pose estimate that is much further from truth.

Table 7.9 indicates that in some cases the modified algorithm produces base amplitude estimates with a larger standard deviation than the original algorithm. For instance, the base amplitude estimates produced by the modified algorithm for tophat 1 have a standard deviation of 0.90 dBsm in the 10°-AVSD experiment and 1.51 dBsm in the 20°-AVSD experiment, compared to standard deviations of 0.32 dBsm and 0.47 dBsm produced by the original algorithm. This is due to correlation between subtype classification and the base amplitude estimate. In particular, the base amplitude that minimizes $Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})$ will change conditional on primitive subtype. For instance, the standard deviation of the base amplitude estimates of tophat 1 condi-

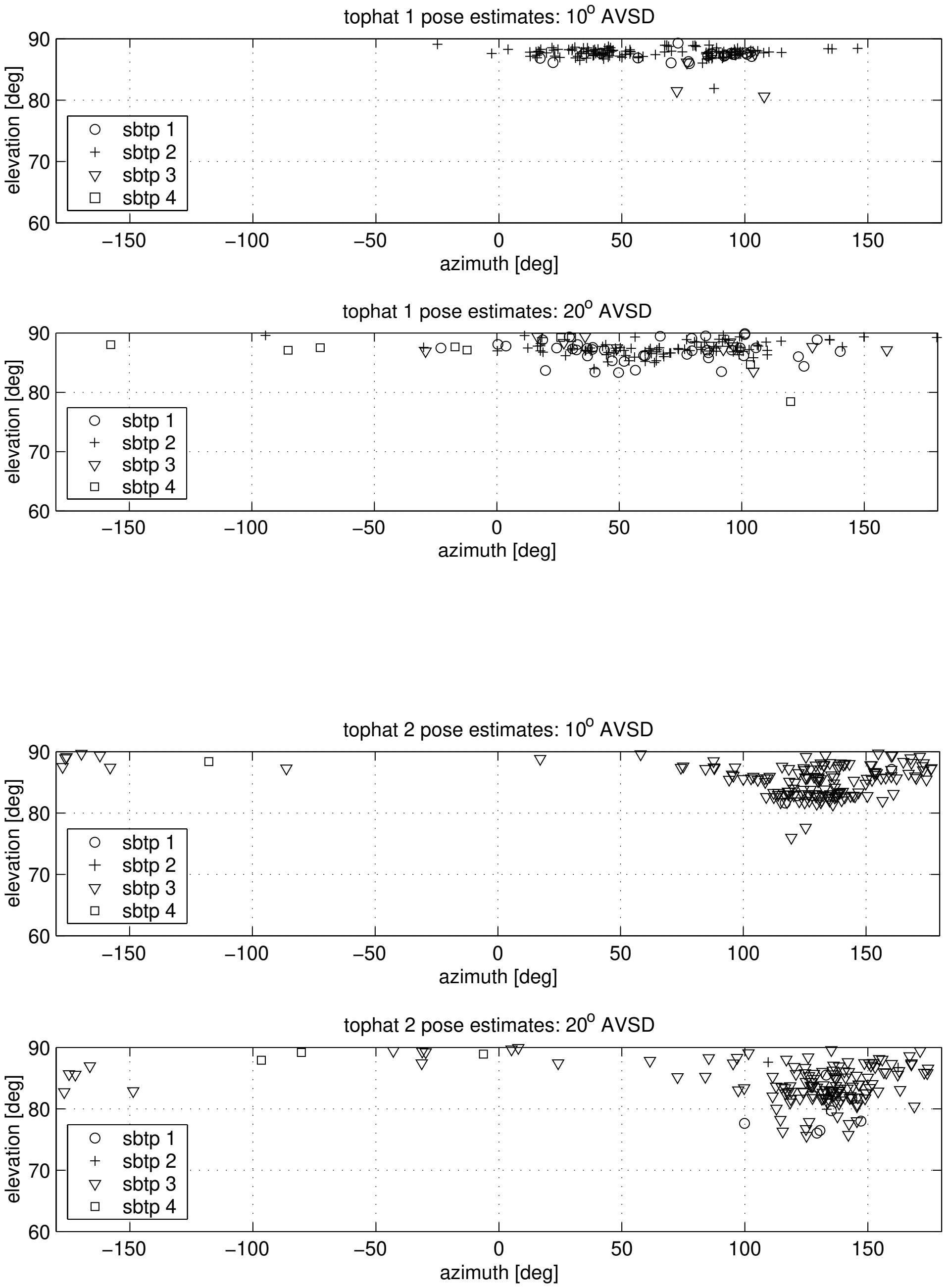


Figure 7.15. Tophat pose estimates produced by the modified algorithm. The true pose of each tophat is 90° elevation, the horizontal line at the top of each plot. The scale on each plot is identical to that of Figure 7.5. These plots also indicate the subtype classification for each estimate.

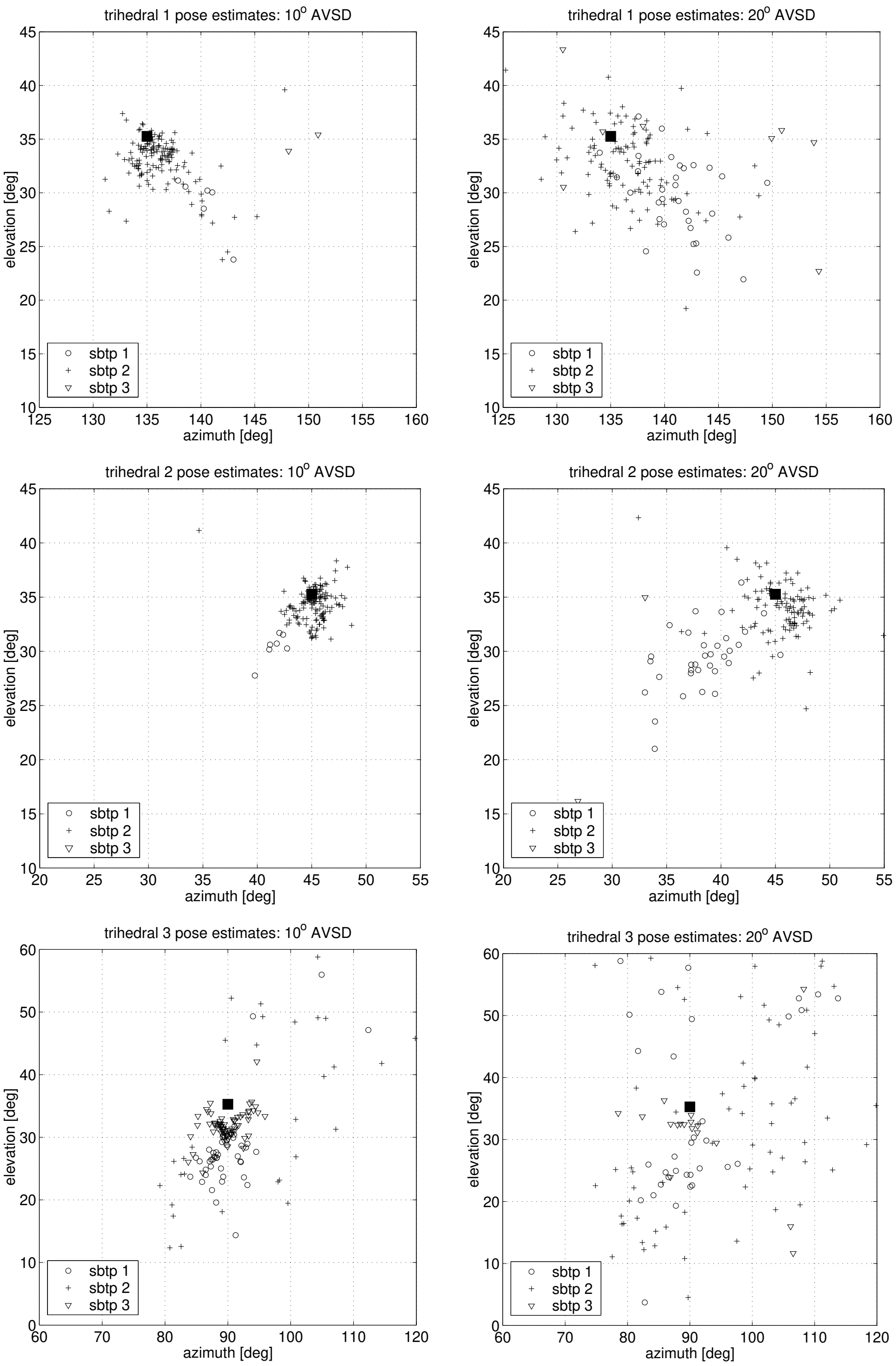


Figure 7.16. Trihedral pose estimates produced by the modified algorithm. The true pose of each trihedral is denoted by the filled square in each plot. The scale on each plot is identical to that of Figure 7.6. These plots also indicate the subtype classification for each estimate.

tional on a classification of subtype 2 is only 0.27 dBsm in the 10°-AVSD experiment and 0.36 dBsm in the 20°-AVSD experiment. This is well below the ensemble values of 0.90 dBsm and 1.51 dBsm.

Figure 7.17 presents a visual depiction of a characteristic set of estimates generated in the 10°-AVSD experiment. In particular, this figure contains four plots, each corresponding to a single run in the 10°-AVSD experiment. Each plot overlays the locations of all primitives produced by the model generation algorithm on an outline of SLICY, viewed from the top. The estimated primitive types are indicated by different symbols, as explained in the figure caption. This figure illustrates many of the effects described above, including the generation of spurious trihedral estimates corresponding to multiple-primitive reflections, the generation of spurious estimates corresponding to the cavity response of tophat 1, and the relatively low discoverability of SLICY's dihedral components. (A more detailed description of these estimates is provided in the caption of Figure 7.17.)

7.6 Distinguishing between Legitimate and Spurious Primitive Estimates

Ideally, a target model generation algorithm would provide not only a set of primitive estimates $\boldsymbol{\theta}_i^{[n]}$, but also some measure of its confidence in each estimate. Such a confidence statistic would be useful for many reasons. For instance, it could be used to guide further modeling by identifying responses that are poorly described using the standard parameterization; such responses could be classified as a special primitive type and described using a different parameterization or in terms of an "all-purpose" set of basis functions. A confidence measure could also be used by a model-based ATR system to aid in the construction and testing of classification hypotheses by providing a measure of the uncertainty associated with various target features.

In this section we describe the construction of a confidence statistic that can be used to distinguish between legitimate and spurious primitive estimates, *i.e.*, between estimates corresponding to actual modeled target components and estimates corresponding to unmodeled response mechanisms. This is motivated by the observation in Sections 7.2.2 and 7.5 that in most runs on SLICY our algorithm tends to produce one or more spurious estimates in an attempt to model multiple-primitive reflections or the hollow interior response of tophat 1. Because most of these spurious estimates are classified as trihedrals, and because SLICY contains three highly observable legitimate trihedral components, we focus on distinguishing between legitimate and spurious trihedral estimates. Furthermore, although an appropriately constructed confidence statistic could be used on its own merits to provide a concise and fine-grained measure of certainty in an estimate, in this section we only consider construction of such a statistic as it relates to the classification of responses as legitimate or spurious. In Chapter 8 we discuss the construction and use of a confidence statistic in a more general context.

In the absence of prior information, a confidence statistic must be based on how well a primitive estimate $\boldsymbol{\theta}_i^{[n]}$ fits the observed data. The expected-log-likelihood $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$

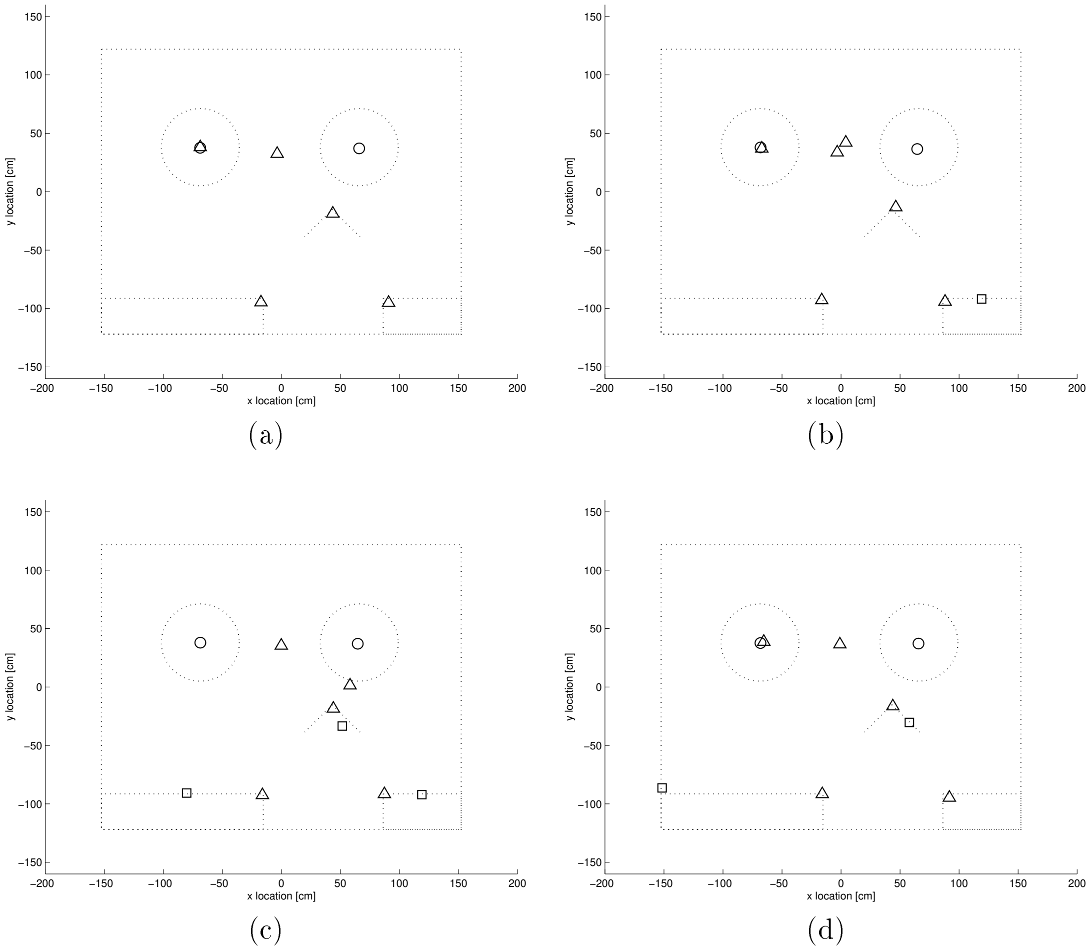


Figure 7.17. Illustration of four SLICY models. The estimated locations of all primitives in four different 10°-AVSD runs are overlayed on an outline of SLICY from the top. The shapes of the overlayed symbols indicate different estimated types: triangles correspond to trihedral estimates, circles correspond to tophat estimates, and squares correspond to dihedral estimates. (None of the models depicted included a cylinder estimate.) Note that in all four runs, the types of SLICY's trihedral and tophat components were correctly estimated, and the locations of these components were very near the true values. The various dihedrals, on the other hand, were discovered in fewer runs: none were discovered in plot (a), one was discovered in plot (b), three were discovered in plot (c), and two were discovered in plot (d). (Note that the location of the left cut-out dihedral in plot (d) was estimated at the edge of that component; this is because the reports responsible for that estimate were nonspecular dihedral returns, *i.e.*, returns generated at viewing angles outside of the dihedral response mainlobe, and which thus appear to emanate from the edge of the primitive as described in Section 3.1 [62].) Note also that all runs contain at least one spurious primitive estimate: for instance, the cavity response from tophat 1 results in a spurious trihedral estimate with x-y location near that of tophat 1 in plots (a), (b), and (d). Similarly, spurious estimates corresponding to the multiple-primitive responses from the two tophats are present in all four plots. (These are located between the two tophats.) Plot (c) also depicts a spurious primitive estimate generated by the multiple-primitive response from trihedral 3 and tophat 2.

provides one measure of this. In particular, if we evaluate the expected-log-likelihood at $\boldsymbol{\theta} = \boldsymbol{\theta}^{[n]}$, we would expect to obtain a less negative value from an estimate that provides a good fit to the data than from an estimate that provides a poor fit to the data. Let us consider the components of $Q(\boldsymbol{\theta}^{[n]}|\boldsymbol{\theta}^{[n]})$ that depend on a single primitive estimate $\boldsymbol{\theta}_i^{[n]}$:

$$\begin{aligned}
Q_i(\boldsymbol{\theta}_i^{[n]}) = &\sum_{k=1}^{K}\sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \Bigg[-\frac{1}{2}\left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i^{[n]})\right)^T \mathbf{R}^{-1}\left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_i^{[n]})\right) \\
&\qquad - \frac{1}{2\sigma_a^2}(a_{k,j} - A(\boldsymbol{\theta}_i^{[n]}, \psi_k, \phi_k))^2 + \log \rho'_{k,j} \Bigg] \\
&+ \sum_{k=1}^{K} \Bigg[\Pr(\delta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \log {P_D}'_{k,i} \\
&\qquad + (1 - \Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) - \Pr(\delta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})) \log(1 - {P_D}'_{k,i}) \Bigg] \\
&+ \sum_{k=1}^{K} \Pr(\beta_{i,k} = 1|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \log \frac{P_b}{1 - P_b}.
\end{aligned} \tag{7.23}$$

This quantity has three conceptually distinct components. In particular, the first two lines of (7.23) penalize the total mismatch between the predicted response of a primitive described by $\boldsymbol{\theta}_i^{[n]}$ and the observed response among the set of reports that correspond to the primitive; the next two lines penalize the mismatch between the predicted and observed detectability of the primitive across the viewsphere; the final line penalizes the attribution of missed detections to obstruction according to the obstruction prior specified in (6.14). For a given set of reports, each of these components will tend to penalize spurious estimates more than legitimate estimates: spurious estimates will likely provide a poorer fit to the observed report measurements, will likely have many missed detections, and will likely attribute many of these missed detections to obstruction. The statistic $Q_i(\boldsymbol{\theta}_i^{[n]})$ thus appears informative for discriminating between legitimate and spurious estimates.

One problem with the use of $Q_i(\boldsymbol{\theta}_i^{[n]})$ to distinguish between legitimate and spurious estimates is that it tends to penalize $\boldsymbol{\theta}_i^{[n]}$ that are based on a large number of reports. In particular, if we define

$$\lambda_T = \sum_{k=1}^{K}\sum_{j=1}^{M_k} \Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}), \tag{7.24}$$

then the first two lines of (7.23) indicate that $Q_i(\boldsymbol{\theta}_i^{[n]})$ will tend to take on more negative

component	10° AVSD		20° AVSD	
	$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ mean	$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ std dev	$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ mean	$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ std dev
trihedral 1	−1.93	0.25	−1.79	0.41
trihedral 2	−2.62	0.48	−2.47	0.89
trihedral 3	−3.53	0.97	−3.78	1.36
all spurious	−7.10	1.21	−5.39	1.33

Table 7.10. Mean and standard deviation of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ for all trihedral estimates.

values as λ_T increases. We thus normalize $Q_i(\boldsymbol{\theta}_i^{[n]})$ by λ_T^{-1} to obtain

$$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]}) = \frac{1}{\lambda_T} Q_i(\boldsymbol{\theta}_i^{[n]}). \tag{7.25}$$

Conceptually, $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$, like $Q_i(\boldsymbol{\theta}_i^{[n]})$, contains three distinct components. The first penalizes the *average* mismatch between the predicted response of $\boldsymbol{\theta}_i^{[n]}$ and the observed report measurements, the second penalizes the *relative* number of views at which there is a significant detection mismatch compared to the number of views at which a measurement is available, and the third penalizes the *relative* number of views at which there is an obstruction penalty compared to the number of views at which a measurement is available. The normalization of the obstruction penalty is intuitively appealing: although attribution of a handful of missed detections to obstruction in the course of explaining a large collection of reports is not generally indicative of a spurious estimate, attribution of a large number of missed detections to obstruction in an attempt to explain a handful of reports almost certainly is.

Figure 7.18 depicts histograms of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ for all trihedral estimates in the 10°-AVSD experiment. The first three histograms are compiled from the estimates produced by trihedrals 1, 2, and 3, respectively; the fourth histogram is compiled from all other trihedral estimates, *i.e.*, all of the spurious trihedrals produced in the 10°-AVSD experiment. Figure 7.19 presents similar histograms for the 20°-AVSD experiment. Figures 7.18 and 7.19 depict a clear difference in the $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ values attained by legitimate and spurious estimates. In particular, the legitimate-estimate histograms and spurious-estimate histograms show relatively little overlap in $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$. Table 7.10 lists the mean and standard deviation for each histogram at each AVSD. Note that although there are differences in the mean and spread of the $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ among the three legitimate primitives, the mean $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ for each of the legitimate primitives is separated by at least one standard deviation (and often several) from the mean spurious value. (Note that at each AVSD, trihedral 3 produces the most negative mean value and displays the greatest spread; this is attributable to the particular difficulties associated with the estimate of this trihedral's parameters, as described in Sections 7.2.2 and 7.5.)

The preceding analysis suggests that legitimate and spurious estimates could be distinguished by application of a simple threshold test. In particular, for a specified

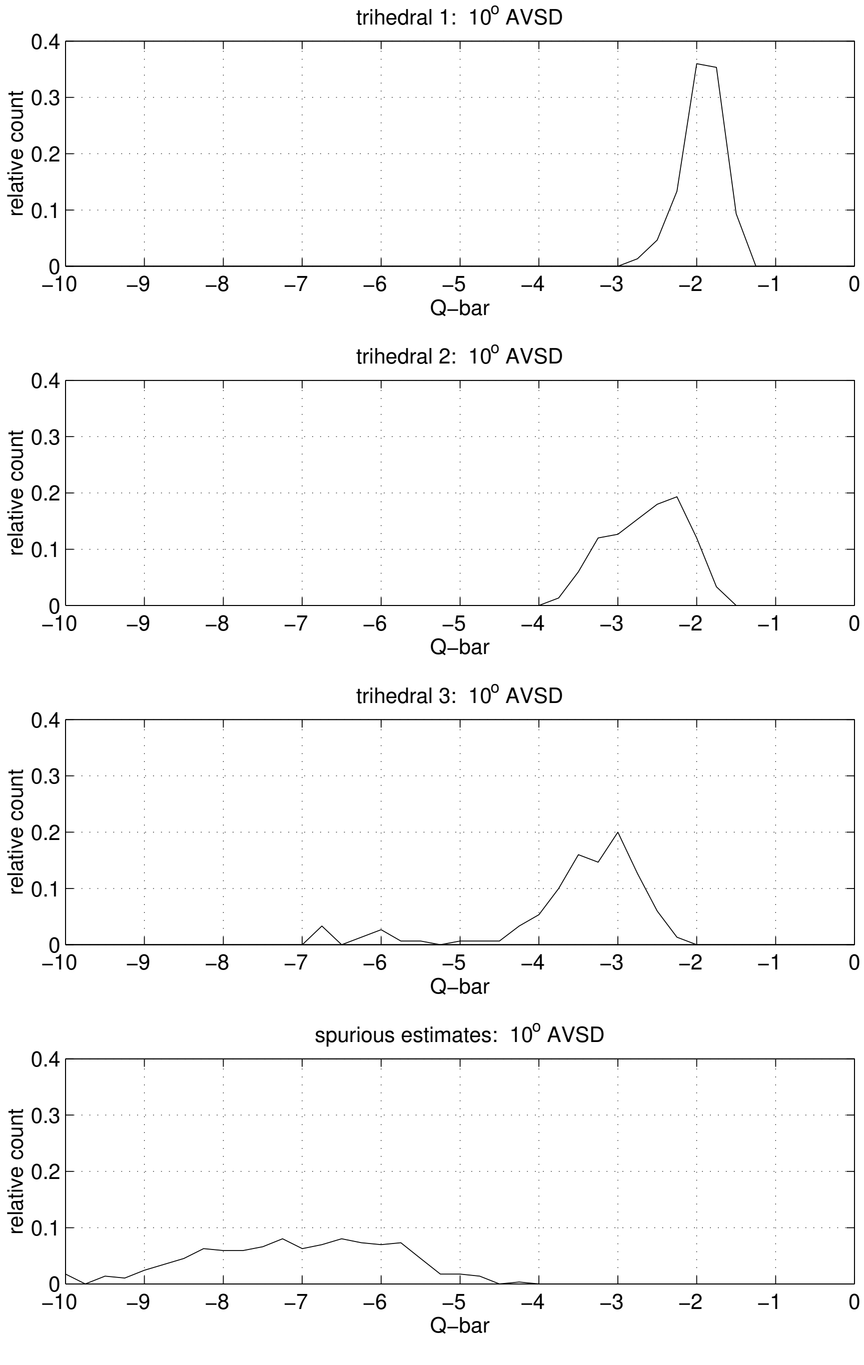


Figure 7.18. Histograms of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ values in the 10°-AVSD experiment. The first three histograms present values attained by estimates of each of the three legitimate primitives; the fourth histogram presents values attained by the spurious estimates.

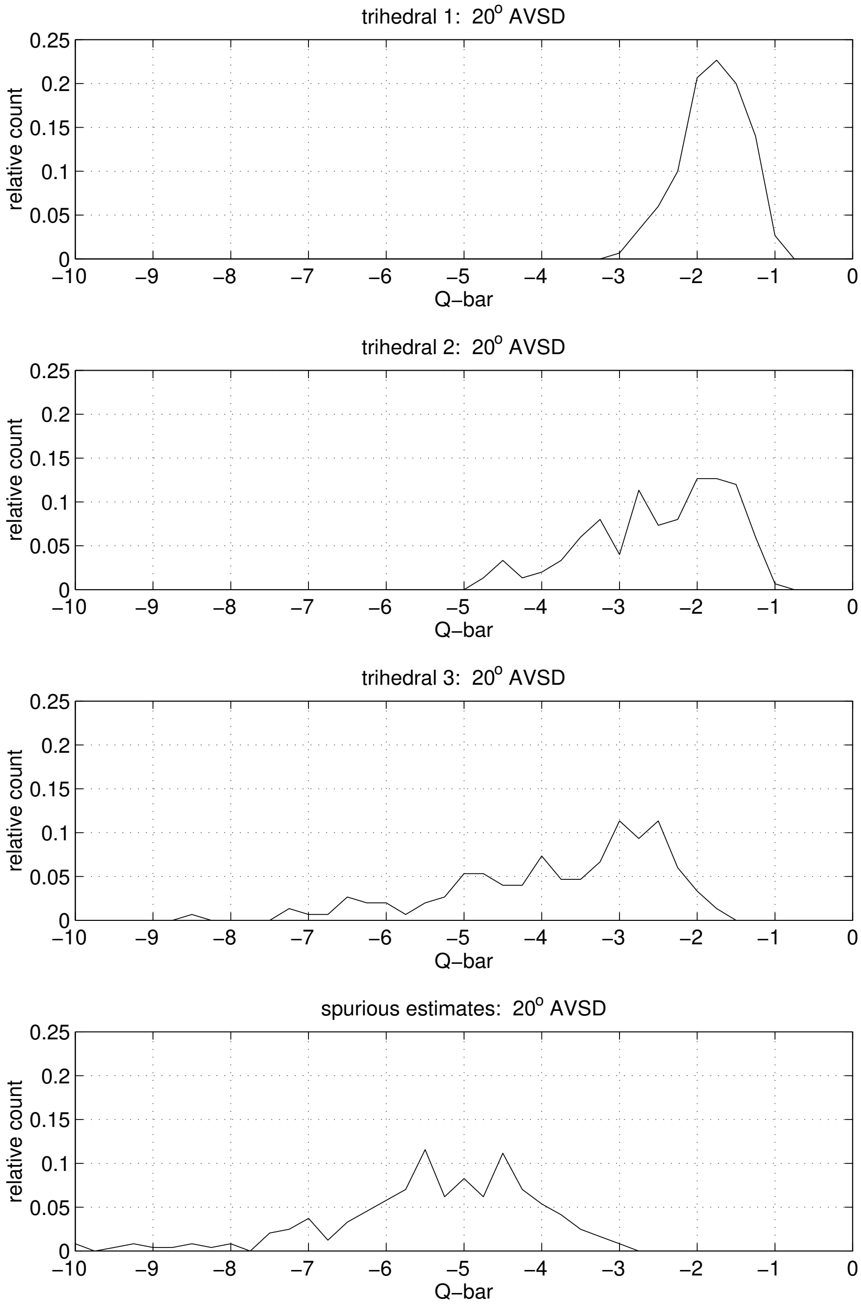


Figure 7.19. Histograms of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ values in the 20°-AVSD experiment. The first three histograms present values attained by estimates of each of the three legitimate primitives; the fourth histogram presents values attained by the spurious estimates.

threshold η_Q, we would apply the following test:

$$\begin{aligned} \bar{Q}_i(\boldsymbol{\theta}_i^{[n]}) \geq \eta_Q : &\quad \text{declare } \boldsymbol{\theta}_i^{[n]} \text{ legitimate;} \\ \bar{Q}_i(\boldsymbol{\theta}_i^{[n]}) < \eta_Q : &\quad \text{declare } \boldsymbol{\theta}_i^{[n]} \text{ spurious.} \end{aligned} \tag{7.26}$$

The threshold η_Q could be chosen by analysis of histograms of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$ values typically attained by legitimate and spurious estimates, such as provided by Figures 7.18 and 7.19. Figure 7.20 depicts two ROC curves [105] obtained by application of (7.26) to the histograms at each AVSD for a range of choices of η_Q. The horizontal axis of each plot represents the fraction of spurious estimates that are incorrectly declared legitimate for a given choice of η_Q; the vertical axis represents the fraction of legitimate estimates that are correctly identified as such for the same choice of η_Q. Each plot indicates points on the curve attained by several different choices of η_Q. Figure 7.20 indicates that near-perfect discrimination between SLICY's legitimate and spurious trihedral estimates is theoretically possible at an AVSD of 10° for appropriate selection of η_Q; performance is worse at an AVSD of 20° because of the lesser separation between the legitimate and spurious histograms.

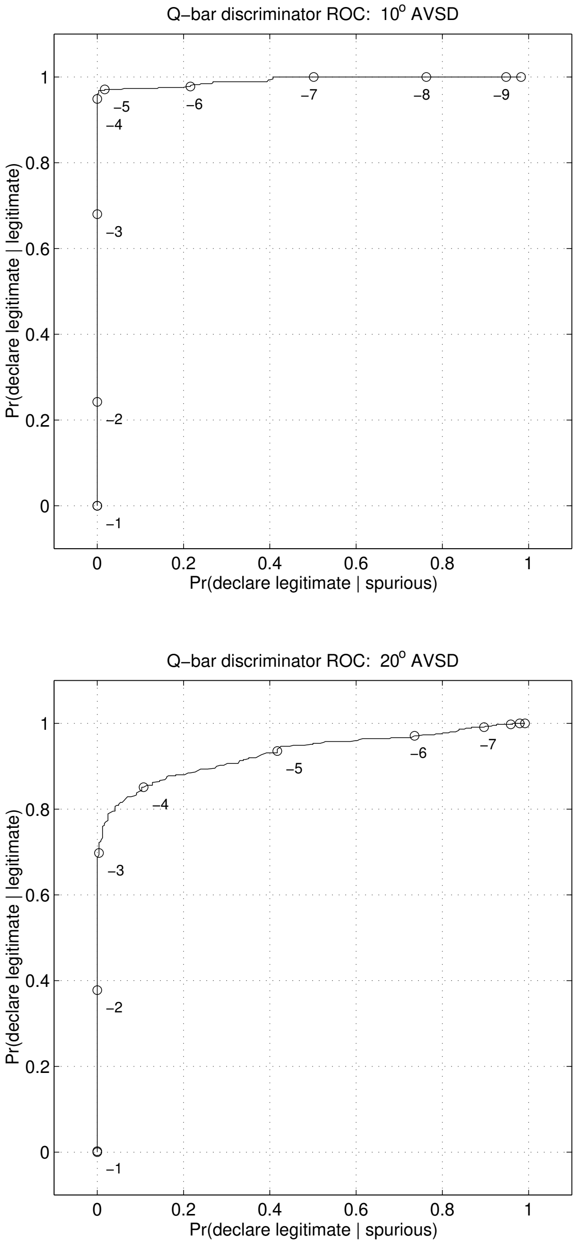


Figure 7.20. ROC curves depicting $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$-based discrimination. The horizontal axis of each plot corresponds to the fraction of spurious estimates correctly classified as spurious for a given choice of $\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$; the vertical axis corresponds to the fraction of true estimates incorrectly classified as spurious. Various values of η_Q are indicated on each plot.

Chapter 8

Contributions and Suggestions for Future Research

THIS thesis makes contributions in the area of target modeling for model-based ATR systems operating on SAR imagery. In this final chapter we summarize the contributions of the thesis and suggest several directions for future research.

8.1 Contributions

The fundamental contribution of this thesis is the development and demonstration of a systematic formalism and flexible framework for the generation of target models from SAR images. In Chapter 3 we described the basic structure of this framework and cast target model generation as a parametric estimation problem whereby we seek a parametric description of a target in terms of its component reflector primitives. We proposed to perform this estimation not directly from the raw set of SAR imagery, but rather from a lower-dimensional summary of that data in the form of a collection of extracted image features. This has two primary benefits: first, it facilitates the specification of a measurement model in terms of the unknown target-primitive identities of the extracted image features, and second, it enables the consideration of model generation from the perspective of data association, enabling the application of the expectation-maximization (EM) method as proposed in Chapter 4. This framework also has the conceptual benefit of allowing an extremely broad range of choices of target parameterization and SAR-image feature extraction.

The major contribution of Chapter 4 is the development an EM-based algorithm to generate target models in the data association framework constructed in Chapter 3. This algorithm treats the image-feature identities as a set of hidden "label parameters" that, although unobservable, provide a conceptual foothold for the estimation of a target model by the EM method. In particular, our algorithm generates a target model by iterating between an expectation (E) step and a maximization (M) step. In the E step the probabilities of each feasible set of label parameters are calculated, conditional on the observed image features and on the target model provided by the previous iteration; in the M step these probabilities are used to estimate a new target model from the data. Because exact calculation of the E-step probabilities is combinatorially difficult, we applied gating to make implementation of this component of the algorithm tractable. We described a specific implementation of the E and M steps of the algorithm for the specific measurement model constructed in Chapter 3. Although some of the details of this implementation depend on the precise form of this measurement model, a similar algorithm could be constructed for any measurement model that relates extracted image features to target parameters via a set of hidden label parameters.

Two further contributions of Chapter 4 are the construction of an initialization technique to produce an estimate with which to seed the EM iteration, and the modification of the basic EM method to enable adaptive selection of model order as the iteration progresses. The initialization procedure we proposed is an agglomerative clustering technique for grouping image features based on their backprojection into $\mathbb{R}^3$. (An extension to the basic initialization technique proposed in Chapter 4 was presented in Chapter 7.) The model-order adjustment stage takes place after the M step in every iteration. It operates by examining the empirical evidence for the existence of each hypothesized target primitive as provided by the E-step probabilities. Hypothesized primitives with little empirical evidence for existence—*i.e.*, hypothesized primitives that explain few image features—are removed.

In Chapter 5 we presented results demonstrating the basic utility of our algorithm and general framework. We characterized the fundamental differences in observability between different types of primitives, and illustrated how these differences impact the estimation of primitives of each basic type. We characterized the performance of our algorithm on a set of single-primitive targets and then demonstrated that when faced with more complex multiple-primitive targets posing a more difficult data association problem, the algorithm can successfully disambiguate the various responses to produce estimates of similar quality to those obtained from the single-primitive targets.

The fundamental contribution of Chapter 6 is the expansion of the basic framework to accommodate partially obstructed primitives, an important consideration in the modeling of real targets. We proposed to accommodate obstruction in a manner that was extremely well-tailored to the existing framework. In particular, we introduced an additional set of hidden variables that enabled the consideration of obstruction effects in the EM framework. These variables took the form of binary obstruction indicators for each primitive in each image. We described how these variables could be incorporated into our algorithm with a few simple modifications. The E step of the modified algorithm entails computation of the probabilities that each of these additional hidden variables takes on either of its possible values. The M step then utilizes these probabilities to obtain a new target model estimate. We presented results demonstrating the effectiveness of this method.

In Chapter 7 we proposed a further expansion of our framework to enable explicit modeling of primitive noncanonicity, another important phenomenon exhibited by real targets. We proposed a general methodology whereby specific classes of noncanonicity are accommodated by augmentation to the target parameterization. Although such an approach would in general result in a significant increase in the computational requirements of the algorithm, we proposed a fundamental modification to the M step that limits the computational impact of such an augmentation. In particular, we proposed to separate the estimation of primitive type from that of all other parameters. In the modified M step, type classification is achieved by maximizing a type-dependent empirical likelihood function for a data-dependent summary statistic. Estimation of all other primitive parameters is then achieved conditional on this type classification. This has the effect of making unnecessary the maximization over all primitive types' augmented parameters to obtain a single primitive's parameter estimate. We implemented our general approach in the context of a specific class of noncanonicities important for proper modeling of SLICY, a benchmark target of considerable interest to the model-based ATR community. We demonstrated that the modified algorithm is capable of successfully accommodating many of the noncanonical effects displayed by SLICY. A final contribution of Chapter 7 was the outline of a method for distinguishing between estimates corresponding to legitimate target primitives and those corresponding to unmodeled response mechanisms.

■ 8.2 Suggestions for Future Research

The research presented in this thesis suggests numerous directions for further investigation. In this final section we describe several such directions. We list them in rough order of increasing conceptual challenge.

Super-Resolution of Scatterers by the Feature Extractor

One shortcoming of the feature extraction technique described in Section 3.2 and utilized throughout this thesis is its inability to resolve closely located target primitives. (This presented a particular difficulty in the estimation of the parameters of one of SLICY's trihedral components in Chapter 7.) Modification of the feature extractor to enable super-resolution of distinct scatterers would correct this shortcoming. There are a wide variety of techniques for super-resolution of features in SAR image data [7, 15, 24, 51, 86, 93]. The techniques of [86, 93], described in Section 2.4.2, are designed not only to super-resolve scattering centers but also to provide estimates of various non-locational attributes of those scattering centers. Utilization of one of these techniques would enable more robust extraction of SAR image features. As long as the chosen feature-extraction technique could be characterized statistically in the manner of Section 5.1.2, replacement of the existing feature extractor with a more sophisticated technique could be accomplished with no fundamental modification to the existing algorithm.

Modeling Dependences between Report Components

One of the basic assumptions used to facilitate construction of a measurement model in Section 3.3 was the conditional independence of individual report components given $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$. This is an unrealistic assumption. For instance, high-amplitude image features—

i.e., features with a high signal-to-clutter ratio—will tend to be extracted more accurately than low-amplitude image features and will thus tend to exhibit smaller location covariances and type confusions. Additionally, as described in Section 3.1, the reports corresponding to double-bounce returns from trihedrals exhibit large location errors because they do not emanate from the trihedral apex; this, coupled with the fact that such reports will usually be classified as even-bounce reflections, implies that there is a dependence between the location and polarimetric-signature-type components for trihedral reports. A more realistic measurement model would describe dependences such as these. The use of such a measurement model would not significantly complicate the algorithm: as long as the measurement-model pdf could be easily evaluated for the components of any report $\mathbf{Z}_{k,j}$, the E and M steps of the algorithm could proceed with no fundamental changes. The primary difficulty facing implementation of such a modification is the task of learning these dependences reliably. Although complete specification of the measurement model in its current form requires determination of only a handful of parameters (summarized in Table 3.2), complete specification of a measurement model in which report components are interdependent would likely require determination of a much wider set of parameters.

Construction of a Finer-Grained Confidence Statistic for Primitive Estimates

In Section 7.6 we introduced the concept of a confidence statistic that would quantify the degree of certainty in the estimates produced for each target primitive. We also outlined a method for distinguishing between legitimate and spurious primitive estimates. This method essentially provides a binary confidence statistic representing a hard-threshold decision about the validity of each estimate. A more general and informative approach would be to provide a continuously varying score statistic. Although the expected-log-likelihood function or the related $\bar{Q}_i$ statistic of Section 7.6 each provide a continuously varying measure of the goodness-of-fit of an estimate to the data, these statistics are not truly informative in their own right. They are informative only in a relative sense: the value of $\bar{Q}_i$ attained by any primitive estimate is meaningful only in comparison to a value attained by other estimates, or in comparison to a threshold determined from a representative set of such estimates. A better confidence statistic would be informative not only in a relative sense, but also in an absolute sense. For instance, one absolutely informative confidence statistic would be a score between 0 and 1 that represents the probability of an estimate being legitimate.

The histograms of $\bar{Q}_i$ presented in Section 7.6 could be used as the basis for construction of a more informative confidence statistic. In particular, these histograms can be considered empirical approximations of the pdfs of $\bar{Q}_i$ under the competing hypotheses that an estimate is legitimate or spurious. In this interpretation, assuming that any estimate is *a priori* as likely to be spurious as legitimate, any value of $\bar{Q}_i$ attained by any primitive estimate could be used to specify a posterior probability of legitimacy:

$$P_l(\bar{Q}_i) = \frac{p_l(\bar{Q}_i)}{p_l(\bar{Q}_i) + p_s(\bar{Q}_i)}, \tag{8.1}$$

where $p_l(\bar{Q}_i)$ is the pdf for the $\bar{Q}_i$ statistic under the hypothesis that primitive i is legitimate, and $p_s(\bar{Q}_i)$ is the pdf for the $\bar{Q}_i$ statistic under the alternative hypothesis that primitive i is spurious. The validity of $P_l(\bar{Q}_i)$ as a true measure of legitimacy would depend on the accuracy of the empirical pdfs $p_l(\bar{Q}_i)$ and $p_s(\bar{Q}_i)$ in capturing the variation in $\bar{Q}_i$ for legitimate and spurious estimates. If pdfs were available for a statistic other than $\bar{Q}_i$, or if a different statistic were found to be more informative in distinguishing between legitimate and spurious estimates, that statistic could be used as the basis for P_l in (8.1).

Expansion of the Set of Extracted SAR Image Features

The feature extraction technique we use throughout this thesis is based on a simple peak extractor, as described in Section 3.2. It provides a relatively coarse-grained description of image features in terms of their locations, amplitudes, and basic polarimetric signature types. While this feature extractor is sufficient to provide a basic description of SAR images, an extraction technique that provides a more fine-grained representation of image features would provide more information about the primitives that generate those features, and could thus aid in the construction of an accurate target model. For instance, our feature extraction technique neglects relevant information such as the GTD frequency dependence of scatterer responses [54, 85, 86], scatterer anisotropy across the imaging aperture [57,58], and polarimetry information beyond that provided by the existing binary type-classification parameter [92,104]. Development of a feature extractor to provide this information, or utilization of an existing SAR image feature extraction technique that provides some of this information (such as those of [86, 93]) could aid the generation of target models. In particular, additional image-feature information could be used not only to improve the quality of estimates of existing target parameters, but also to aid in the estimation of additional parameters in an augmented target parameterization. (We discuss expansion of the existing target parameterization shortly.)

Conceptually, any feature extractor that provides measurements that can be related to individual target primitives in terms of the set of hidden label parameters could be used in the existing framework. One extremely general approach to feature extraction would be to characterize SAR image chips or phase histories in terms of a library of appropriately chosen basis functions in the range/cross-range domain (for images) or the frequency/azimuth domain (for phase histories).[1] A feature extractor operating in

[1] In fact, such an approach underlies the super-resolution algorithms of [86, 93]: although these techniques describe scatterers in terms of a handful of physical parameters similar to those provided by our feature extractor, this is accomplished by decomposing a phase history into a sum of complex frequency/azimuth basis functions.

this way would provide not merely a single scalar amplitude parameter $a_{k,j}$, but instead a vector of basis-function coefficients $\mathbf{a}_{k,j}$. (This approach is similar to those of [14, 70], although these efforts focus on the characterization of one-dimensional range profiles rather than SAR images.) Such an approach would provide a more flexible means for describing observed SAR image features by enabling accurate characterization of responses not succinctly or naturally described by a handful of physically motivated parameters. For instance, such an approach could enable more natural description of resonant responses generated by periodic symmetries in target structure [62, 70], dispersive responses generated by cavities or similar structures [14, 62, 70], or non-localized responses generated by target components that produce strong returns over multiple resolution cells. Challenges to the implementation of such an approach would be the construction of an appropriate library of basis functions and the development of a measurement model relating $\mathbf{a}_{k,j}$ to $\boldsymbol{\theta}$ for that library.

Augmentation of Primitive Descriptor Vectors to Include Additional Physical Attributes

The target parameterization developed in Section 3.1 relies on a small but descriptive set of parameters to describe the physical attributes and predicted scattering responses of target primitives. In Chapter 7 we proposed an augmentation to this basic parameterization to include a discrete parameter indexing a set of "subtype" scattering responses for each basic primitive type. Although either parameterization enables compact and informative representation of a variety of primitives and their responses, it could be expanded to include additional physically based primitive descriptors. (We discuss augmentation or replacement of the existing parameterization to include a more general class of descriptors, *i.e.*, descriptors that do not necessarily correspond to physical primitive attributes, shortly.) For instance, as described in Section 3.1, our chosen parameterization does not include descriptors for primitive dimension (other than the dimension-related base amplitude), although this quantity provides important physical information and is useful in predicting the scattering response of a primitive. In Chapter 7 we describe other potential augmentations to the primitive parameterization. As described in Chapter 7, the primary effect of the inclusion of additional primitive descriptors on the algorithm would be an increase in complexity of the M step. In particular, maximization over additional primitive descriptors will likely impose a significant additional computational burden. (Investigation into maximization techniques other than those used here is thus another potential area of research.)

A fundamental question associated with the augmentation or modification of the primitive descriptor vector is how to choose descriptors that are not only informative but also robustly estimable. Inclusion of parameters that cannot be reliably estimated from extracted SAR image features would offer no benefit to a model-based ATR system and would have the drawback of complicating the implementation of a target model generation algorithm. Augmentation of the primitive descriptor vector to include additional physical attributes should be guided by investigation of the observability of potential descriptors in the primitive response. Development of a principled approach to such an analysis represents a conceptual challenge.

Proper Modeling of Multiple-Primitive Responses

A fundamental assumption (perhaps *the* fundamental assumption) of our framework is that any image feature can be uniquely associated with a single target component, and that any target component produces at most a single feature in any image. This assumption dictates the basic structure of our measurement model and algorithm. In particular, the E step of the algorithm calculates the probability of correspondence of each image feature to each hypothesized target primitive under the assumption that there is a unique set of correspondences; the M step then provides a parametric target model by utilizing these correspondence probabilities to weight each measurement. Modification of this assumption to allow a single image feature to be attributed to multiple primitives, or to allow a primitive to generate multiple image features (*e.g.*, not only a "canonical" feature but also a multiple-primitive feature through interaction with another target component) could require fundamental modifications to the basic structure of our framework.

Conceptually, one approach to accommodating multiple-primitive responses would be to expand the definition of the label parameters $\boldsymbol{\lambda}$ to recognize such responses. For instance, we could modify (3.15) to write

$$\lambda_{k,j} = \begin{cases} i, & \text{if report } \mathbf{Z}_{k,j} \text{ corresponds to target primitive } i, \\ 0, & \text{if report } \mathbf{Z}_{k,j} \text{ is spurious (corresponds to no primitive),} \\ -1 & \text{if report } \mathbf{Z}_{k,j} \text{ is a multiple-primitive response.} \end{cases} \tag{8.2}$$

For reports for which $\lambda_{k,j} = -1$, we could introduce a new hidden variable $\boldsymbol{\mu}_{k,j}$ describing the origin of the report:

$$\boldsymbol{\mu}_{k,j} = \begin{cases} \begin{bmatrix} i_1 & \cdots & i_n \end{bmatrix} & \text{if report } \mathbf{Z}_{k,j} \text{ corresponds to primitives } i_1, \ldots, i_n, \\ 0, & \text{if report } \mathbf{Z}_{k,j} \text{ is not a multiple-primitive response.} \end{cases} \tag{8.3}$$

This would require a change in the measurement model to enable description of multiple-primitive reports in terms of the primitives indicated by $\boldsymbol{\mu}_{k,j}$. Such a modification would probably be nontrivial—even predicting the locations of multiple-primitive responses in the slant plane could require high-level geometrical reasoning. Even if such a modification to the measurement model could be achieved in a straightforward fashion, this modification would have important ramifications on the E and M steps of the algorithm. The E step would require calculation of the probabilities of $\boldsymbol{\mu}_{k,j}$ taking on each of its possible values; the M step would entail simultaneous maximization over all $\boldsymbol{\theta}_i$ due to the primitive coupling introduced by the modeling of multiple-primitive reflections.

Assuming the E step were tractable, one possibility for circumventing the M-step primitive coupling would be simply to neglect the informative value of the multiple-primitive reports in forming primitive parameter estimates. In other words, we could simply ignore all reports for which $\lambda_{k,j} = -1$ in the implementation of the M step, and determine each $\boldsymbol{\theta}_i$ using only $\Pr(\lambda_{k,j} = i|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ as in the existing algorithm. This would be an approximation to the optimal approach that might serve to make consideration of multiple-primitive reports tractable.

Development of Empirical Response-Based Primitive Descriptors

Most of the attention in this thesis has been given to primitive descriptors that are directly tied to physical attributes of the primitives they represent. For instance, pose and location each describe a basic physical primitive attribute. Other potential descriptors we have considered, such as additional dimensions or the angles between primitive components, are also physically based. A different approach to the construction of a primitive parameterization framework is to employ descriptors that are more naturally viewed not as physical attributes, but as scattering-response attributes. The subtype approach to dealing with noncanonicity introduced in Chapter 7 essentially represents a step in this direction: although the trihedral and tophat subtype responses considered in that chapter were intended to represent the effects of specific geometrical perturbations on the scattering responses of those primitives, such a physical basis was not essential to the specification of subtype responses.

We can imagine specifying the response of any primitive in terms of a superposition of basis functions (possibly unique to each primitive type) on the viewsphere. A primitive descriptor vector could then be constructed to represent the primitive in terms of the attributes of the basis functions that best describe its response. In particular, given a dictionary $\mathcal{D}_{\theta_i^t}$ of basis functions associated with a primitive of type θ_i^t, we could represent the overall scattering response $S(\psi, \phi)$ of that primitive as

$$S(\psi,\phi) = \sum_{j=1}^{n} \gamma_j s_{m_j}(\psi,\phi), \qquad s_{m_j}(\psi,\phi) \in \mathcal{D}_{\theta_i^t}, \tag{8.4}$$

i.e., as a scaled superposition of elements chosen from the dictionary $\mathcal{D}_{\theta_i^t}$. In this framework the descriptors $\theta_i^{\mathbf{p}}$ and θ_i^a would be replaced by a set of descriptors including the set of indices $\{m_j\}_{j=1}^n$ and the set of coefficients $\{\gamma_j\}_{j=1}^n$.[2]

Implementation of the approach just described involves two important conceptual challenges. First is the enforcement of sparsity in the representation of any primitive response, *i.e.*, the limitation of n to be a small integer. This is desirable both in consideration of the basic model-based ATR paradigm and in order to enforce some regularity in the responses from hypothesized primitives. Sparsity could be enforced in a number of ways. For instance, we could fix n to be a previously specified small integer and use matching pursuit [67, 70] to obtain the superposition of n basis elements that minimizes the ℓ^2 mismatch between the predicted and observed primitive scattering

[2] The basic canonical parameterization actually represents a special case of this approach in which each $\mathcal{D}_{\theta_i^t}$ consists of rotations of a single basis function $S_{\theta_i^t}(\cdot)$ on the viewsphere, n is required to equal 1, $\theta_i^{\mathbf{p}}$ selects a particular rotation of $S_{\theta_i^t}(\cdot)$, and θ_i^a scales this basis function.

responses. Alternatively, we could utilize high-resolution pursuit [47], a decomposition technique similar to matching pursuit but which is more sensitive to local structure in the selection of its basis elements. We could also use basis pursuit [17] to obtain a sparse representation of the form of (8.4) by minimization of an objective function that penalizes both the ℓ^2 mismatch between the observed and predicted primitive responses and the ℓ^1 sum of the scaling coefficients.

The second, more fundamental conceptual challenge associated with this approach is the construction of an appropriate dictionary $\mathcal{D}_{\theta_i^t}$ for each primitive type. An effective dictionary would enable accurate modeling of the kinds of responses observed from each primitive type and would simultaneously encourage sparsity by containing basis functions that in some sense capture the primary modes of variation in these responses. There are a number of general techniques geared to the selection of a low-dimensional basis for representation of a family of signals, ranging from pursuit techniques [17, 44, 47, 67] to principal components analysis [52, 55] to other so-called "best basis" selection techniques [20, 90, 101]. In general, however, the selection of an appropriate dictionary for representation of a class of signals is difficult. Construction of response dictionaries for the different primitive types, or construction of a "super-dictionary" capable of succinctly modeling the response from any primitive type, would be a fundamental step in the representation of targets for model-based ATR.

Appendix A

Construction of Primitive Scattering Functional Models

As indicated in Section 3.2, our models for the RCS responses from the four basic primitive types are based on PO and constructed from the components in Table 2.1. Central to these models are the log-shaping functions $S_{\theta_i^t}(.)$ that, along with base amplitude and polarimetric scaling factors, enable us to predict the RCS of a primitive of type θ_i^t as in (3.5), reproduced here for convenience:

$$A(\boldsymbol{\theta}_i, \psi_k, \phi_k) = \theta_i^a + S_{\theta_i^t}(\psi'_{i,k}, \phi'_{i,k}) + c_{\text{pol}}. \tag{A.1}$$

Here we detail the construction of the $S_{\theta_i^t}(.)$ and the other components of (A.1) for each primitive type.

As described in Sections 2.3.2 and 3.1, a single primitive might have not only a primary reflection mechanism but also one or more secondary reflection mechanisms. For instance, a trihedral's primary reflection mechanism is a triple-bounce response, by which incoming rays are backscattered after being reflected in turn from all three component plates; its secondary reflection mechanisms are the single- and double-bounce responses observed at oblique angles, where incoming rays are backscattered after being reflected from only one or two component plates. Figure A.1 depicts the contribution of individual reflection mechanisms to the overall trihedral response: the contributing components include a triple-bounce reflection mechanism (Figure A.1a), three double-bounce reflection mechanisms produced by pairs of the trihedral's plates (Figure A.1b), and three single-bounce reflection mechanisms produced by individual plates (Figure A.1c).[1] These responses are summed to yield the overall response in Figure A.1d. (Azimuth and elevation are defined as in Figure 2.10; the trihedral dimension used to generate the responses of Figure A.1 is $a = 28.0$ cm.)

We construct the $S_{\theta_i^t}(.)$ for the trihedral, tophat, dihedral, and cylinder as sums of individual response mechanisms tabulated in Table 2.1. In particular, we take each

[1] Although the three individual double-bounce components of Figure A.1b are identical except for unique Euler angle orientations, as are the three single-bounce components of Figure A.1c, the projection of azimuth and elevation coordinates from the viewsphere onto a rectangular array distorts each response differently.

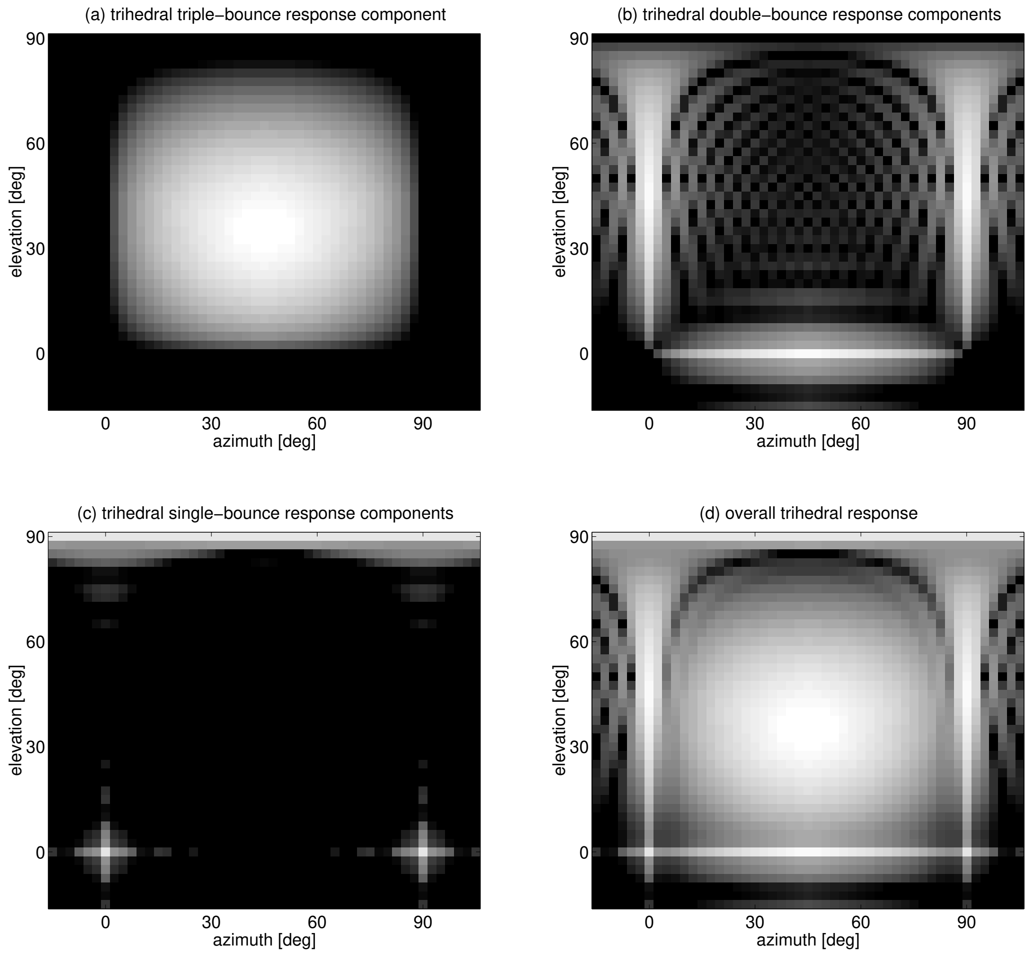


Figure A.1. Construction of trihedral response model from multiple reflection mechanisms. The overall trihedral response (d, lower right) is a combination of several distinct reflection mechanisms: a triple-bounce reflection mechanism (a, upper left), three properly rotated double-bounce reflection mechanisms (b, upper right) and three properly rotated single-bounce mechanisms (c, lower right). Although an observed response would be a coherent sum of these responses, taking the relative phases between the components into account, our models are formed as a noncoherent sum of the responses due to the complication the introduction of primitive dimensions would introduce into the scattering models.

primitive	components of $S_{\theta_i^t}(\psi'_{i,k}, \phi'_{i,k})$	relevant dimensions	θ_i^a
trihedral	one trihedral triple-bounce response three dihedral double-bounce responses three rectangular-plate responses	a	$10\log_{10}\left(\frac{3k_0^2a^4}{\pi}\right)$
tophat	one tophat double-bounce response one cylinder-shaft response	h	$10\log_{10}\left(2k_0rh^2\right)$
dihedral	one dihedral double-bounce response two rectangular-plate responses	a, b	$10\log_{10}\left(\frac{2k_0^2a^2b^2}{\pi}\right)$
cylinder	one cylinder-shaft response	h	$10\log_{10}\left(k_0rh^2\right)$

Table A.1. Primitive scattering model composition.

$S_{\theta_i^t}(\cdot)$ to be a noncoherent sum of one or more of the components of Table 2.1, with any primitive dimensions relevant to the determination of the shaping functions of these components fixed at empirically chosen values (see Chapter 5). The component reflection mechanisms used to construct the $S_{\theta_i^t}(\cdot)$ are listed in Table A.1, along with the primitive dimensions (see Figure 2.10) that affect the component shaping functions.[2]

Construction of the $S_{\theta_i^t}(\cdot)$ in this way embodies two approximations: first, it neglects the dependence of the shaping functions of Table 2.1 on primitive dimensions, and second, it neglects the relative phases between the component reflection mechanisms (dependent on primitive dimensions and relative viewing angle) that will affect the coherent combination of the primary and secondary reflection mechanisms to dictate the overall response. As indicated in Section 3.2, these approximations are made to avoid the algorithmic complication the introduction of primitive dimension would create. The effects of these approximations are examined in Chapter 5.

Each $S_{\theta_i^t}(\cdot)$ is scaled to give a maximum value of 0 dBsm, so that θ_i^a in (A.1) is taken for each primitive type to be the value indicated in Table A.1. The remaining term in (A.1) is c_{pol}, the effective gain achieved by using multiple polarimetric channels, *i.e.*, by taking RCS to be the squared-magnitude of a polarimetric vector $\mathbf{g}$ of the type in (3.7). It can be seen from Table 2.2 that any canonical odd- or even-bounce reflection mechanism will give $\|\mathbf{g}\|^2 = 2$, so c_{pol} is a constant that adds an effective gain of 3 dBsm (*i.e.*, $20\log_{10} 2$) for any primitive type.

[2]We do not include circular-plate components in the overall tophat and cylinder shaping functions. This is motivated by the narrow mainlobe of the circular-plate responses, the large angular separation between these primitives' primary response mechanisms and the circular-plate specular directions, and the relatively computationally intensive Bessel term in the circular-plate response (see Table 2.1).

Appendix B

Derivation of the Expected-Log-Likelihood Function

We require an explicit expression for $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]})$ in order to describe the E step of the EM method as presented in Section 4.2. The general form of the E step is given in (4.5), reproduced here for convenience:

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) &= E\left[\log p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta}) \,\middle|\, \mathbf{Z}, \boldsymbol{\theta}^{[n-1]}\right] \\ &= \sum_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \Big[\log p(\boldsymbol{\lambda}, \mathbf{Z}|\boldsymbol{\theta})\Big] p(\boldsymbol{\lambda}|\mathbf{Z}, \boldsymbol{\theta}^{[n-1]}). \end{aligned} \tag{B.1}$$

Utilizing the independence assumptions of Section 3.3, we can rewrite (B.1) as

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) &= \sum_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \left[\sum_{k=1}^{K} \log p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta})\right] \cdot p(\boldsymbol{\lambda}_k|\mathbf{Z}, \boldsymbol{\theta}^{[n]}) \cdot p(\boldsymbol{\lambda} \setminus \boldsymbol{\lambda}_k|\mathbf{Z}, \boldsymbol{\theta}^{[n]}) \\ &= \sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}\in\boldsymbol{\Lambda}} \Big[\log p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta})\Big] \cdot p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \cdot p(\boldsymbol{\lambda} \setminus \boldsymbol{\lambda}_k|\mathbf{Z} \setminus \mathbf{Z}_k, \boldsymbol{\theta}^{[n]}), \end{aligned} \tag{B.2}$$

where the notation "$\boldsymbol{\lambda} \setminus \boldsymbol{\lambda}_k$" is shorthand for "$\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_{k-1}, \boldsymbol{\lambda}_{k+1}, \ldots, \boldsymbol{\lambda}_K$." (An analogous definition applies to "$\mathbf{Z} \setminus \mathbf{Z}_k$.") It then follows that

$$\begin{aligned} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_{k=1}^{K} &\left[\sum_{\boldsymbol{\lambda}_k\in\boldsymbol{\Lambda}_k} \Big[\log p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta})\Big] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})\right] \\ &\cdot \left[\sum_{\boldsymbol{\lambda}\setminus\boldsymbol{\lambda}_k\in\boldsymbol{\Lambda}\setminus\boldsymbol{\Lambda}_k} p(\boldsymbol{\lambda} \setminus \boldsymbol{\lambda}_k|\mathbf{Z} \setminus \mathbf{Z}_k, \boldsymbol{\theta}^{[n]})\right], \end{aligned} \tag{B.3}$$

because $p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]})$ is constant with respect to $\boldsymbol{\lambda} \setminus \boldsymbol{\lambda}_k$. The final summation in (B.3) clearly evaluates to 1, thus

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = \sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k\in\boldsymbol{\Lambda}_k} \Big[\log p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta})\Big] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}). \tag{B.4}$$

Replacing $\log p(\boldsymbol{\lambda}_k, \mathbf{Z}_k|\boldsymbol{\theta})$ in the above expression with its value according to (3.42) yields, after some rearrangement of terms,

$$\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) = &\sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} \left[-\gamma_{FA} V - \log(M_k!) + \sum_{i=1}^{N} \log\left(1 - P'_{D_{k,i}}\right) \right] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \\
&+ \sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} \sum_{j:\lambda_{k,j}=0} \Big[\log p_{FA}(a_{k,j}) \Big] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \\
&+ \sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} \left[F_k \log\left(\frac{\gamma_{FA}}{2}\right) \right] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \\
&+ \sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} \sum_{j:\lambda_{k,j}\neq 0} \Bigg[\log \frac{P'_{D_{k,\lambda_{k,j}}}}{1 - P'_{D_{k,\lambda_{k,j}}}} - \log 2\pi (\det \mathbf{R})^{1/2} - \log \sigma_a \sqrt{2\pi} \\
&\qquad - \frac{1}{2} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) \right)^T \mathbf{R}^{-1} \left(\mathbf{x}_{k,j} - \boldsymbol{\pi}_k(\boldsymbol{\theta}_{\lambda_{k,j}}) \right) \\
&\qquad + \log \rho'_{k,j} - \frac{1}{2\sigma_a^2} \left(a_{k,j} - A(\boldsymbol{\theta}_{\lambda_{k,j}}, \psi_k, \phi_k) \right)^2 \Bigg] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}).
\end{aligned} \tag{B.5}$$

We will consider the terms of (B.5) line by line. Note that all terms in the first line of this equation are constant given $\mathbf{Z}_k$, so the first line of (B.5) evaluates to

$$\sum_{k=1}^{K} \left[-\gamma_{FA} V - \log(M_k!) + \sum_{i=1}^{N} \log\left(1 - P'_{D_{k,i}}\right) \right]. \tag{B.6}$$

The second line can be rewritten as

$$\sum_{k=1}^{K} \sum_{j=1}^{M_k} \Big[\log p_{FA}(a_{k,j}) \Big] \Pr(\lambda_{k,j} = 0|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}). \tag{B.7}$$

To evaluate the next lines, let us define an indicator function for each $i = 0, 1, \ldots, N$:

$$I_{\{\lambda_{k,j}=i\}} = \begin{cases} 1, & \lambda_{k,j} = i, \\ 0, & \lambda_{k,j} \neq i. \end{cases} \tag{B.8}$$

Then the third line of (B.5) can be written as

$$\sum_{k=1}^{K} \sum_{\boldsymbol{\lambda}_k \in \boldsymbol{\Lambda}_k} \left[\sum_{j=1}^{M_k} I_{\{\lambda_{k,j}=0\}} \log\left(\frac{\gamma_{FA}}{2}\right) \right] p(\boldsymbol{\lambda}_k|\mathbf{Z}_k, \boldsymbol{\theta}^{[n]}), \tag{B.9}$$

which evaluates to

$$\sum_{k=1}^{K}\sum_{j=1}^{M_k}\left[\log\left(\frac{\gamma_{FA}}{2}\right)\right]\Pr(\lambda_{k,j}=0|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}). \tag{B.10}$$

We can use the indicator functions of (B.8) similarly to evaluate the remaining three lines of (B.5) (*i.e.*, the final summation) by noting that for any function $\zeta(\lambda_{k,j})$ we have

$$\begin{aligned}\sum_{k=1}^{K}\sum_{\boldsymbol{\lambda}_k\in\boldsymbol{\Lambda}_k}\sum_{j:\lambda_{k,j}\neq 0}\zeta(\lambda_{k,j})p(\boldsymbol{\lambda}_k|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}) &= \sum_{k=1}^{K}\sum_{\boldsymbol{\lambda}_k\in\boldsymbol{\Lambda}_k}\left[\sum_{i=1}^{N}\sum_{j=1}^{M_k}I_{\{\lambda_{k,j}=i\}}\right]\zeta(\lambda_{k,j})p(\boldsymbol{\lambda}_k|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]})\\ &= \sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{M_k}\zeta(i)\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}),\end{aligned} \tag{B.11}$$

so that the final three lines of (B.5) evaluate to

$$\begin{aligned}\sum_{k=1}^{K}\sum_{i=1}^{N}\sum_{j=1}^{M_k}\Bigg[&\log\frac{P_{D}{}'_{k,i}}{1-P_{D}{}'_{k,i}}-\log 2\pi(\det\mathbf{R})^{1/2}-\log\sigma_a\sqrt{2\pi}\\ &-\frac{1}{2}\left(\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)^T\mathbf{R}^{-1}\left(\mathbf{x}_{k,j}-\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)\right)\\ &+\log\rho'_{k,j}-\frac{1}{2\sigma_a^2}\left(a_{k,j}-A(\boldsymbol{\theta}_i,\psi_k,\phi_k)\right)^2\Bigg]\Pr(\lambda_{k,j}=i|\mathbf{Z}_k,\boldsymbol{\theta}^{[n]}).\end{aligned} \tag{B.12}$$

Combining (B.6), (B.7), (B.10), and (B.12), we can now write (B.5) as

$$\begin{aligned}Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{[n]}) &= \sum_{i=1}^{N}Q_i(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})+C_K\\ &= \sum_{i=1}^{N}\sum_{k=1}^{K}Q_{i,k}(\boldsymbol{\theta}_i|\boldsymbol{\theta}_i^{[n]})+C_K,\end{aligned} \tag{B.13}$$

where

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

and

$$C_K = \sum_{k=1}^{K} \left[-\gamma_{FA} V - \log(M_k!) + \sum_{j=1}^{M_k} \left[\log p_{FA}(a_{k,j}) + \log \left(\frac{\gamma_{FA}}{2} \right) \right] \Pr(\lambda_{k,j} = 0 | \mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \right.$$
$$\left. - \sum_{j=1}^{M_k} \sum_{i=1}^{N} \left(\log 2\pi (\det \mathbf{R})^{1/2} + \log \sigma_a \sqrt{2\pi} \right) \Pr(\lambda_{k,j} = i | \mathbf{Z}_k, \boldsymbol{\theta}^{[n]}) \right]. \tag{B.15}$$

Note that each $Q_{i,k}(\boldsymbol{\theta}_i | \boldsymbol{\theta}_i^{[n]})$ depends only on a single primitive and the reports in a single image, and that C_K is constant with respect to $\boldsymbol{\theta}$.

Appendix C

Construction of Trihedral and Tophat Subtype Responses

In Chapter 7 we utilize a set of trihedral and tophat subtypes intended to model some of the variation in these primitives' geometries associated with a specific type of non-canonicity. In this appendix we detail the construction of the scattering responses for these subtypes.

■ C.1 Trihedral Subtype Responses

The three trihedral subtype responses are depicted in Figure 7.11. As indicated in Section 7.4.4, these responses are heuristic approximations to the types of responses produced by trihedrals whose geometries differ from that of the canonical square-plate trihedral as depicted in Figure 7.9. In particular, compared to the canonical trihedral, subtype 2 will produce stronger responses at viewing angles near its extended axis, and subtype 3 will produce stronger responses at viewing angles near its extended base. GO predicts that the response of any trihedral will depend on the projected area of the region producing a triple-bounce reflection. In particular, if the projected area of this triple-bounce region at a given viewing angle is $A_{\text{tb}}(\psi', \phi')$, then GO predicts a trihedral RCS of [92]

$$\sigma_{\text{tri}} = \frac{4}{\pi} k_0^2 A_{\text{tb}}^2(\psi', \phi'). \tag{C.1}$$

For the canonical square-plate trihedral depicted in Figure 2.10, $A_{\text{tb}}(\psi', \phi')$ can be expressed in closed form:

$$A_{\text{tb}}(\psi', \phi') = \begin{cases} \frac{a^2}{2}\left(\frac{4 l_i m_i}{n_i}\right), & m_i \leq \frac{n_i}{2}, \\ \\ \frac{a^2}{2}\left(l_i\left(4 - \frac{n_i}{m_i}\right)\right), & m_i > \frac{n_i}{2}, \end{cases} \tag{C.2}$$

where a is the square-plate length and where l_i, m_i, and n_i are the direction cosines between the viewing angle (ψ', ϕ') and the trihedral axes, ordered so that $l_i \leq m_i \leq n_i$:

$$(l_i, m_i, n_i) = \text{sort}_{\min\to\max}(\sin\psi',\, \cos\psi' \sin\phi',\, \cos\psi' \cos\phi'). \tag{C.3}$$

Note that

$$l_i^2 + m_i^2 + n_i^2 = 1, \tag{C.4}$$

and that $A_{\rm tb}(\psi', \phi')$ (and thus $\sigma_{\rm tri}$) will be maximized when $l_i = m_i = n_i$.

Unfortunately, in the general case in which a trihedral response is produced by three plates of varying shapes or sizes, there is no closed-form expression for $A_{\rm tb}(\psi', \phi')$. Exact determination of the triple-bounce area will generally require numerical integration [94]. To mimic the effects of geometrical perturbations of the type depicted in Figure 7.9 on the trihedral response, we thus rely on a heuristic approximation in which the trihedral response is changed by a modification to (C.3). In particular, to construct the trihedral subtype responses of Figure 7.11 we independently scale each element of (C.3), taking

$$(\tilde{l}_i, \tilde{m}_i, \tilde{n}_i) = \mathrm{sort}_{\min\to\max}\left(\frac{\alpha_1}{\alpha_0'}\sin\psi',\ \frac{\alpha_2}{\alpha_0'}\cos\psi'\sin\phi',\ \frac{\alpha_3}{\alpha_0'}\cos\psi'\cos\phi'\right), \tag{C.5}$$

where α_1, α_2, and α_3 are arbitrary positive coefficients and where α_0' is a normalization coefficient:

$$\alpha_0' = (\alpha_1^2 \sin^2\psi' + \alpha_2^2 \cos^2\psi' \sin^2\phi' + \alpha_3^2 \cos^2\psi' \cos^2\phi')^{-1/2}. \tag{C.6}$$

We then use $\tilde{l}_i, \tilde{m}_i$, and $\tilde{n}_i$ to calculate $A_{\rm tb}(\psi', \phi')$ as in (C.2), which in turn allows calculation of the trihedral response as in (C.1). By selecting appropriate values for α_1, α_2, and α_3 we can perturb the canonical trihedral response so that $A_{\rm tb}(\psi', \phi')$ and $\sigma_{\rm tri}$ are maximized at a viewing angle closer to one of the trihedral axes, as is the case for a trihedral of subtype 2, or at a viewing angle closer to one of the trihedral bases, as is the case for a trihedral of subtype 3. We obtain the response for subtype 2 using

$$\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix} = \begin{bmatrix}0.5\\ 1\\ 1\end{bmatrix}, \tag{C.7}$$

and the response for subtype 3 using

$$\begin{bmatrix}\alpha_1\\ \alpha_2\\ \alpha_3\end{bmatrix} = \begin{bmatrix}0.6\\ 0.6\\ 1\end{bmatrix}. \tag{C.8}$$

■ C.2 Tophat Subtype Responses

As with the trihedral, the GO approximation to the tophat response relies on the determination of the projected area of a multiple-bounce region—in this case, the projected area $A_{\rm db}(\psi', \phi')$ of the tophat base and cylinder-shaft region generating a double-bounce reflection. In particular, GO predicts that given this double-bounce area, the RCS of a tophat with cylinder radius r will be [12]

$$\sigma_{\rm top} = \frac{k_0}{4r} A_{\rm db}^2(\psi', \phi'). \tag{C.9}$$

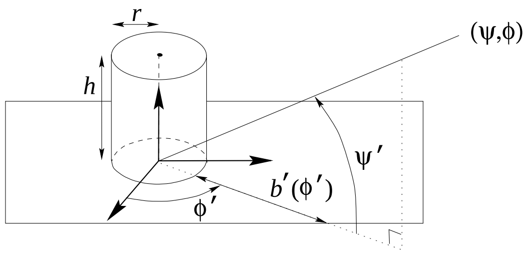


Figure C.1. Illustration of general rectangular-base-tophat geometry.

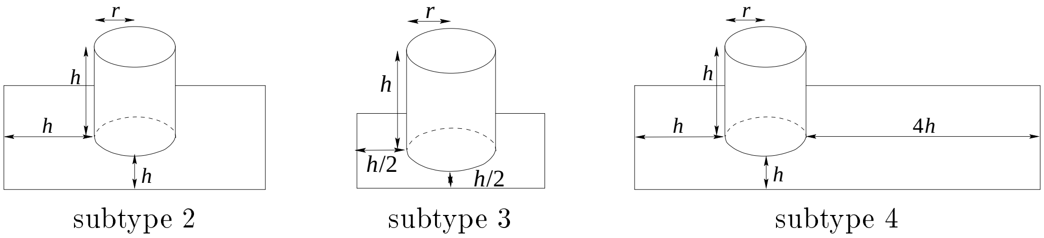


Figure C.2. Dimensions used to construct square- and rectangular-base tophat subtype responses.

For a circular-base tophat with cylinder radius r, base radius $b + r$, and height h, the projected double-bounce area can be shown to be approximately [12]

$$A_{\mathrm{db}}(\psi', \phi') = \begin{cases} 4rb \sin \psi', & 0 \leq \psi' < \arctan\left(\frac{h}{b}\right), \\ 4rh \cos \psi', & \arctan\left(\frac{h}{b}\right) \leq \psi' \leq \frac{\pi}{2}. \end{cases} \tag{C.10}$$

When $b = h$, substitution of (C.10) into (C.9) gives the canonical tophat response as listed in Table 2.1. For tophats of arbitrary base shape, however, determination of $A_{\mathrm{db}}(\psi', \phi')$ is problematic and in the general case requires numerical integration [12]. We thus rely on an approximation to $A_{\mathrm{db}}(\psi', \phi')$ to form the subtype responses for the four tophat subtypes depicted in Figure 7.12. In particular, consider the tophat geometry depicted in Figure C.1. This tophat has height h, cylinder radius r, and a rectangular base whose extent from the cylinder at any viewing azimuth ϕ' is denoted by $b'(\phi')$, as pictured. We assume that the double-bounce region of such a tophat can be approximated by the corresponding double-bounce region of a *circular*-base tophat of cylinder radius r, base radius $b'(\phi') + r$, and height h. We thus model

$$A_{\mathrm{db}}(\psi', \phi') = \begin{cases} 4rb'(\phi') \sin \psi', & 0 \leq \psi' < \arctan\left(\frac{h}{b'(\phi')}\right), \\ 4rh \cos \psi', & \arctan\left(\frac{h}{b'(\phi')}\right) \leq \psi' \leq \frac{\pi}{2}, \end{cases} \tag{C.11}$$

which then allows RCS to be modeled as in (C.9). We construct the latter three tophat subtype responses depicted in Figure 7.14 in this way, using square- and rectangular-base dimensions as pictured in Figure C.2.

Appendix D

Summary of Notation

Symbol	Definition
General Notation	
$\lvert\,\cdot\,\rvert$	absolute value
$\lVert\,\cdot\,\rVert$	2-norm
$\lVert\,\cdot\,\rVert_{\mathbf{Q}}$	$\mathbf{Q}$-norm (for positive-definite $\mathbf{Q}$), *i.e.*, $\lVert\mathbf{x}\rVert_{\mathbf{Q}} = (\mathbf{x}^T\mathbf{Q}\mathbf{x})^{1/2}$
$\binom{\cdot}{\cdot}$	combination operator, *i.e.*, $\binom{n}{m} = \frac{n!}{m!(n-m)!}$
$\backslash$	set difference
$\hat{\cdot}$	estimator of specified quantity
$*$	convolution operator, *i.e.*, $x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau)h(t-\tau)\,d\tau$
#	cardinality
$\emptyset$	empty set
det	determinant
$E(\cdot)$	expectation
$p(\cdot)$	probability density function or probability mass function
$\Pr(\cdot)$	probability of specified event
$\mathbb{R}$	real numbers
$\mathbb{R}_+$	non-negative real numbers
$\mathbb{R}^n$	space of real n-vectors
$\mathrm{sinc}(\cdot)$	sinc function, *i.e.*, $\mathrm{sinc}(x) = \frac{\sin x}{x}$
$W_{[u,v]}(\cdot)$	1-D unit-amplitude window function extending from u to v
$\mathbb{Z}_+$	non-negative integers
Radar and Propagation Phenomenology	
c	speed of light
$\mathcal{E}(\cdot)$ or $\mathcal{E}[\cdot]$	continuous or discrete measurements of field strength
f	frequency
f_c	carrier or center frequency
$g(\cdot,\cdot)$	2-D complex reflectivity density function

Symbol	Definition
Radar and Propagation Phenomenology, continued	
k_0	wavenumber, *i.e.*, $k_0 = \frac{2\pi f}{c}$
T	transmit-pulse duration
α	PO- or GTD-predicted frequency-dependence exponent
δ_r	range resolution
δ_{xr}	cross-range resolution
η_c	chirp rate
σ	radar cross section
σ_{tr}	radar cross section for transmit/receive polarization tr
ϕ	absolute viewing azimuth
ϕ'	relative viewing azimuth
ψ	absolute viewing elevation/depression
ψ'	relative viewing elevation/depression
Target Parameterization and Scattering Models	
$A(\boldsymbol{\theta}_i, \psi, \phi)$	predicted log-RCS of primitive i at elevation ψ, azimuth ϕ
c_{pol}	effective multiple-polarimetric-channel gain
$\mathbf{H}_k$	2×3 ground-to-slant-plane projection matrix for image k
N	target model order
n_t	number of discrete primitive types
$S_t(\psi', \phi')$	PO-based log-shaping function describing relative amplitude of a primitive of type t at relative viewing elevation ψ' and relative viewing azimuth ϕ'
$\boldsymbol{\theta}$	target parameter vector
$\boldsymbol{\theta}_i$	parameter vector for primitive i
θ_i^a	base amplitude of primitive i
$\boldsymbol{\theta}_i^{\mathbf{d}}$	all parameters for primitive i other than type and location
$\theta_i^{\mathbf{p}}$	Euler-angle pose of primitive i
θ_i^r	radius of curvature of primitive i
θ_i^t	type of primitive i
$\theta_i^{\mathbf{x}}$	location of primitive i
$\boldsymbol{\pi}_k(\boldsymbol{\theta}_i)$	predicted location of primitive i in image k
ϕ_k	center viewing azimuth of image k
$\phi'_{i,k}$	relative viewing azimuth of primitive i in image k
ψ_k	center viewing elevation of image k
$\psi'_{i,k}$	relative viewing elevation of primitive i in image k

Symbol	Definition
SAR-Image Features	
$a_{k,j}$	extracted log-magnitude amplitude of report j in image k
F_k	number of false alarms in image k
K	number of SAR images in the data set
M_k	number of reports in image k
$t_{k,j}$	binary polarimetric-signature type classification of report j in image k
$\mathbf{x}_{k,j}$	extracted 2-D location of report j in image k
$\mathbf{Z}_{k,j}$	feature vector for report j in image k
$\mathbf{Z}_k$	feature vector for all reports in image k
$\mathbf{Z}$	feature vector for reports in all images
Measurement Model	
$p_{FA}(\cdot)$	false-alarm-amplitude probability density function
$P_D(\cdot)$	amplitude-dependent probability-of-detection function
${P_D}'_{k,i}$	predicted probability of detection of primitive i in image k
$\mathbf{R}$	extracted-location covariance
V	sensor volume
γ_{FA}	false alarm rate
$\delta_{i,k}$	detection indicator for primitive i in image k
$\boldsymbol{\delta}_k$	vector of detection indicators for all primitives in image k
$\lambda_{k,j}$	label parameter for report j in image k
$\boldsymbol{\lambda}_k$	vector of label parameters for all reports in image k
$\boldsymbol{\lambda}$	vector of label parameters for all reports in all images
$\boldsymbol{\Lambda}_k$	space of all possible $\boldsymbol{\lambda}_k$
$\boldsymbol{\Lambda}$	space of all possible $\boldsymbol{\lambda}$
$\boldsymbol{\rho}$	extracted-type confusion matrix
$\rho'_{k,j}$	$(\boldsymbol{\theta}_{\lambda_{k,j}}, t_{k,j})$ element of $\boldsymbol{\rho}$
σ_a	extracted-amplitude standard deviation
Algorithm	
$a_{\min}$	minimum report amplitude for clustering in initialization
$n_{\min}$	minimum number of reports in cluster for initialization
P_{disc}	probability of primitive discovery
P_{id}	discovery-conditional probability of correct primitive type identification

Symbol	Definition
Algorithm, continued	
$Q(\boldsymbol{\theta}\|\boldsymbol{\theta}^{[n]})$	expected-log-likelihood function
$Q_i(\boldsymbol{\theta}_i\|\boldsymbol{\theta}_i^{[n]})$	primitive-i-dependent components of $Q(\boldsymbol{\theta}\|\boldsymbol{\theta}^{[n]})$
$Q_{i,k}(\boldsymbol{\theta}_i\|\boldsymbol{\theta}_i^{[n]})$	image-k-dependent components of $Q_i(\boldsymbol{\theta}_i\|\boldsymbol{\theta}_i^{[n]})$
r_{gate}	E-step gate radius
η	chi-squared-statistic dissimilarity threshold in initialization
$\boldsymbol{\theta}^{[n]}$	estimate of $\boldsymbol{\theta}$ from iteration n
$\boldsymbol{\theta}_i^{[n]}$	estimate of $\boldsymbol{\theta}_i$ from iteration n
$\theta_i^{\cdot\,[n]}$	estimate of $\theta_i^{\cdot}$ from iteration n
Obstruction	
P_b	prior probability of obstruction of any primitive in any image
$\beta_{i,k}$	obstruction indicator for primitive i in image k
$\boldsymbol{\beta}_k$	vector of obstruction indicators for all primitives in image k
$\boldsymbol{\beta}$	vector of obstruction indicators for all primitives in all images
$\mathbf{B}_k$	space of all possible $\boldsymbol{\beta}_k$
$\mathbf{B}$	space of all possible $\boldsymbol{\beta}$
Noncanonicity	
f_d	fraction of images in data set with detections of given primitive
f_t	fraction of detections of given primitive with odd-bounce $t_{k,j}$
$\mathbf{M}$	detection-viewing-angle orientation matrix
$n_s(t)$	number of subtypes for a primitive of type t
$Q_i(\boldsymbol{\theta}_i^{[n]})$	components of $Q(\boldsymbol{\theta}^{[n]}\|\boldsymbol{\theta}^{[n]})$ that depend on $\boldsymbol{\theta}_i^{[n]}$
$\bar{Q}_i(\boldsymbol{\theta}_i^{[n]})$	primitive-estimate-legitimacy discrimination statistic
$\mathbf{s}(\hat{\boldsymbol{\lambda}}, \mathbf{Z})$	summary statistic for type classification
v_n	nth largest eigenvalue of $\mathbf{M}$
θ_i^s	subtype of primitive i
$\hat{\lambda}_{k,j}$	label-parameter estimate for report j in image k
$\hat{\boldsymbol{\lambda}}$	vector of label-parameter estimates for all reports in all images

Appendix E

Glossary of Acronyms

Acronym	Definition
ATR	automatic target recognition
AVSD	average viewsphere sampling density
CAD	computer-aided design
cdf	cumulative distribution function
dBsm	decibel-scale square meters
DFT	discrete Fourier transform
ECM	expectation-conditional-maximization
EM	expectation-maximization
FM	frequency modulation
HH	horizontal-transmit, horizontal-receive polarization
HV	horizontal-transmit, vertical-receive polarization
GLRT	generalized likelihood-ratio test
GO	geometrical optics
GTD	geometrical theory of diffraction
LLSE	linear least-squares error
ML	maximum likelihood
MRF	Markov random field
MSTAR	moving and stationary target acquisition and recognition
pdf	probability density function
pmf	probability mass function
PO	physical optics

Acronym	Definition
PWDE	Parzen-window density estimator
RCS	radar cross section
rmse	root-mean-squared error
ROC	receiver-operating characteristic
ROI	region of interest
SAR	synthetic aperture radar
SBR	shooting and bouncing rays
SLICY	Sandia Laboratories implementation of cylinders
TLS	total least-squares
VH	vertical-transmit, horizontal-receive polarization
VV	vertical-transmit, vertical-receive polarization

Bibliography

[1] M. R. Anderberg. *Cluster Analysis for Applications.* Academic Press, New York, 1973.

[2] T. W. Anderson and M. A. Stephens. Tests for randomness of directions against equatorial and bimodal alternatives. *Biometrika*, 59(3):613–621, December 1972.

[3] W. C. Anderson. Consequences of nonorthogonality on the scattering properties of dihedral reflectors. *IEEE Transactions on Antennas and Propagation*, AP-35(10):1154–1159, October 1987.

[4] J. Baldauf, A.-W. Lee, L. Lin, A.-K. Jeng, S. M. Scarborough, and C. L. Yu. High frequency scattering from trihedral corner reflectors and other benchmark targets: SBR versus experiment. *IEEE Transactions on Antennas and Propagation*, 39(9):1345–1351, September 1991.

[5] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association.* Academic Press, Boston, 1988.

[6] R. Barakat. Weak-scatterer generalization of the K-density function with application to laser scattering in atmospheric turbulence. *Journal of the Optical Society of America A*, 3(4):401–409, April 1986.

[7] G. R. Benitz. High-definition vector imaging for synthetic aperture radar. In *Conference Record of the Asilomar Conference on Signals, Systems & Computers*, volume 2, pages 1204–1209, Pacific Grove, CA, October 1997.

[8] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995.

[9] B. Bhanu. Automatic target recognition: State of the art survey. *IEEE Transactions on Aerospace and Electronic Systems*, 22(4):364–379, July 1986.

[10] B. Bhanu, D. E. Dudgeon, E. G. Zelnio, A. Rosenfeld, D. Casasent, and I. S. Reed. Guest editorial introduction to the special issue on automatic target detection and recognition. *IEEE Transactions on Image Processing*, 6(1):1–6, January 1997.

[11] R. A. Boyles. On the convergence of the EM algorithm. *Journal of the Royal Statistical Society B*, 45(1):47–50, 1983.

[12] C. G. Brown and K. Sarabandi. Geometrical and physical optics solutions for the backscatter radar cross section of tophat calibration targets. In *Proceedings of the International Geoscience and Remote Sensing Symposium*, pages 2715–2717, Honolulu, HI, July 2000.

[13] R. Buderi. *The Invention That Changed the World: How a Small Group of Radar Pioneers Won the Second World War and Launched a Technological Revolution.* Simon & Schuster, New York, 1996.

[14] L. Carin, L. B. Felsen, D. R. Kralj, H. S. Oh, W. C. Lee, and S. U. Pillai. Wave-oriented signal processing of dispersive time-domain scattering data. *IEEE Transactions on Antennas and Propagation*, 45(4):592–600, April 1997.

[15] W. G. Carrara, R. S. Goodman, and R. M. Majewski. *Spotlight Synthetic Aperture Radar: Signal Processing Algorithms.* Artech House, Norwood, MA, 1995.

[16] R. Carrière and R. L. Moses. High resolution radar modeling using a modified Prony estimator. *IEEE Transactions on Antennas and Propagation*, 40(1):13–18, January 1992.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, August 1998.

[18] D.-M. Chiang. *Parametric Signal Processing Techniques for Model Mismatch and Mixed Parameter Estimation.* PhD thesis, The Ohio State University, Columbus, OH, 1996.

[19] D. T. Clemens and D. W. Jacobs. Space and time bounds on indexing 3-D models from 2-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(10):1007–1017, October 1991.

[20] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, March 1992.

[21] J. W. Crispin, Jr. and K. M. Siegel, editors. *Methods of Radar Cross-Section Analysis.* Academic Press, New York, 1968.

[22] J. Cui and W. Freeden. Equidistribution on the sphere. *SIAM Journal on Scientific Computing*, 18(2):595–609, March 1997.

[23] J. C. Curlander and R. N. McDonough. *Synthetic Aperture Radar: Systems and Signal Processing.* John Wiley & Sons, New York, 1991.

[24] S. R. DeGraaf. SAR imaging via modern 2-D spectral estimation methods. *IEEE Transactions on Image Processing*, 7(5):729–761, May 1998.

[25] DEMACO, Inc., Champaign, IL. *User's Manual for XPatchES*, 1991. Version 2.4.

[26] DEMACO, Inc., Champaign, IL. *User's Manual for XPatchF*, 1991. Version 2.4.

[27] DEMACO, Inc., Champaign, IL. *User's Manual for XPatchT*, 1991. Version 2.4.

[28] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society B*, 39(1):1–38, 1977.

[29] G. A. Deschamps. Ray techniques in electromagnetics. *Proceedings of the IEEE*, 60(9):1022–1035, September 1972.

[30] L. Devroye. *A Probabilistic Theory of Pattern Recognition.* Number 31 in Applications of Mathematics. Springer-Verlag, New York, 1996.

[31] D. E. Dudgeon and R. T. Lacoss. An overview of automatic target recognition. *Lincoln Laboratory Journal*, 6(1):3–10, Spring 1993.

[32] E. Ertin and L. C. Potter. Polarimetric classification of scattering centers. In E. G. Zelnio and R. J. Douglass, editors, *Algorithms for Synthetic Aperture Radar Imagery III*, volume 2757 of *Proceedings of SPIE*, pages 206–216, Orlando, FL, April 1996.

[33] G. Ferrara, F. Mattia, and F. Posa. Backscattering study on non-orthogonal trihedral corner reflectors. *IEE Proceedings—Microwaves, Antennas and Propagation*, 142(6):441–446, December 1995.

[34] N. I. Fisher, T. Lewis, and B. J. J. Embleton. *Statistical Analysis of Spherical Data.* Cambridge University Press, Cambridge, 1987.

[35] S. C. Fralick and R. W. Scott. Nonparametric Bayes-risk estimation. *IEEE Transactions on Information Theory*, IT-17(4):440–444, July 1971.

[36] N. S. Friedland. Feature characterization in the context of model-based search. In E. G. Zelnio and R. J. Douglass, editors, *Algorithms for Synthetic Aperture Imagery III*, volume 2757 of *Proceedings of SPIE*, pages 253–270, Orlando, FL, April 1996.

[37] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, November 1984.

[38] D. Giuli. Polarization diversity in radars. *Proceedings of the IEEE*, 74(2):245–269, February 1986.

[39] T. Griesser and C. A. Balanis. Backscatter analysis of dihedral corner reflectors using physical optics and the physical theory of diffraction. *IEEE Transactions on Antennas and Propagation*, AP-35(10):1137–1147, October 1987.

[40] P. Hall, S. Sheather, M. C. Jones, and J. S. Marron. On optimal data-based bandwidth selection in kernel density estimation. *Biometrika*, 78(2):263–269, June 1991.

[41] M. Hazlett, D. J. Andersh, S. W. Lee, H. Ling, and C. L. Yu. XPatch: A high-frequency electromagnetic scattering prediction code using shooting and bouncing rays. In W. R. Watkins and D. Clement, editors, *Targets and Backgrounds: Characterization and Representation*, volume 2469 of *Proceedings of SPIE*, pages 266–274, Orlando, FL, April 1995.

[42] F. B. Hildebrand. *Introduction to Numerical Analysis*. McGraw-Hill, New York, 1974.

[43] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[44] P. J. Huber. Projection pursuit. *The Annals of Statistics*, 13(2):435–475, February 1985.

[45] M. P. Hurst and R. Mittra. Scattering center analysis via Prony's method. *IEEE Transactions on Antennas and Propagation*, AP-35(8):986–988, August 1987.

[46] J. R. Huynen. *Phenomenological Theory of Radar Targets*. PhD thesis, Delft University of Technology, Delft, Netherlands, 1970.

[47] S. Jaggi. *Multiscale Geometric Feature Extraction and Object Recognition*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1997.

[48] E. Jakeman and P. N. Pusey. Significance of K distributions in scattering experiments. *Physical Review Letters*, 40(9):546–550, 27 February 1978.

[49] E. Jakeman and R. J. A. Tough. Generalized K distribution: a statistical model for weak scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772, September 1987.

[50] C. V. Jakowatz, Jr. and P. A. Thompson. A new look at spotlight mode synthetic aperture radar as tomography: Imaging 3-D targets. *IEEE Transactions on Image Processing*, 4(5):699–703, May 1995.

[51] 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.

[52] I. T. Jolliffe. *Principal Component Analysis.* Springer-Verlag, New York, 1986.

[53] I. R. Joughin, D. B. Percival, and D. P. Winebrenner. Maximum likelihood estimation of K distribution parameters for SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 31(5):989–999, September 1993.

[54] J. B. Keller. Geometrical theory of diffraction. *Journal of the Optical Society of America*, 52(2):116–130, February 1962.

[55] M. G. Kendall. *Multivariate Analysis.* Hafner Press, New York, 1975.

[56] E. R. Keydel and S. W. Lee. Signature prediction for model-based automatic target recognition. In E. G. Zelnio and R. J. Douglass, editors, *Algorithms for Synthetic Aperture Imagery III*, volume 2757 of *Proceedings of SPIE*, pages 306–317, Orlando, FL, April 1996.

[57] A. J. Kim, S. Dogan, J. W. Fisher III, R. L. Moses, and A. S. Willsky. Attributing scatterer anisotropy for model based ATR. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery VII*, volume 4053 of *Proceedings of SPIE*, pages 176–188, Orlando, FL, April 2000.

[58] A. J. Kim, J. W. Fisher III, A. S. Willsky, and P. A. Viola. Nonparametric estimation of aspect dependence for ATR. In E. G. Zelnio, editor, *Algorithms for Synthetic Aperture Radar Imagery VI*, volume 3721 of *Proceedings of SPIE*, pages 332–342, Orlando, FL, April 1999.

[59] A. Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems.* Number 120 in Applied Mathematical Sciences. Springer, New York, 1996.

[60] E. F. Knott. RCS reduction of dihedral corners. *IEEE Transactions on Antennas and Propagation*, AP-25(5):406–409, May 1977.

[61] E. F. Knott. A progression of high-frequency RCS prediction techniques. *Proceedings of the IEEE*, 73(2):252–264, February 1985.

[62] E. F. Knott, J. F. Shaeffer, and M. T. Tuley. *Radar Cross Section.* Artech House, Norwood, MA, 2nd edition, 1993.

[63] T. Kurien. Issues in the design of practical multitarget tracking algorithms. In Y. Bar-Shalom, editor, *Multitarget-Multisensor Tracking: Advanced Applications*, chapter 3, pages 43–83. Artech House, Norwood, MA, 1990.

[64] K. Lange and R. Carson. EM reconstruction algorithms for emission and transmission tomography. *Journal of Computer Assisted Tomography*, 8(2):306–316, April 1984.

[65] J. S. Lim. *Two-Dimensional Signal and Image Processing.* Prentice-Hall, Englewood Cliffs, NJ, 1990.

[66] H. Ling, R.-C. Chou, and S.-W. Lee. Shooting and bouncing rays: Calculating the RCS of an arbitrarily shaped cavity. *IEEE Transactions on Antennas and Propagation*, 37(2):194–205, February 1989.

[67] S. G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41(12):3397–3415, December 1993.

[68] L. Marple. *Digital Spectral Analysis with Applications*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[69] B. Marsolf, S. Kosanovich, D. Andersh, and J. Hughes. Large scale multidisciplinary electromagnetics computations. In *Proceedings of the 30th AIAA Plasmadynamics and Lasers Conference*, Norfolk, VA, June 1999.

[70] M. McClure and L. Carin. Matching pursuits with a wave-based dictionary. *IEEE Transactions on Signal Processing*, 45(12):2912–2927, December 1997.

[71] G. J. McLachlan and T. Krishnan. *The EM Algorithm and Extensions*. John Wiley & Sons, New York, 1997.

[72] X.-L. Meng and D. B. Rubin. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2):267–278, June 1993.

[73] Model Based Vision Laboratory, Air Force Research Laboratory, Wright-Patterson Air Force Base, Dayton, OH. MSTAR (public) targets. CDROM, 1997.

[74] Model Based Vision Laboratory, Air Force Research Laboratory, Wright-Patterson Air Force Base, Dayton, OH. MSTAR/IU (public) mixed targets. CDROM, 1998.

[75] T. K. Moon. The expectation-maximization algorithm. *IEEE Signal Processing Magazine*, 13(6):47–60, November 1996.

[76] D. C. Munson. A signal processing view of strip-mapping synthetic aperture radar. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(12):2131–2147, December 1989.

[77] D. C. Munson, Jr., J. D. O'Brien, and W. K. Jenkins. A tomographic formulation of spotlight-mode synthetic aperture radar. *Proceedings of the IEEE*, 71(8):917–925, August 1983.

[78] D. C. Munson, Jr. and J. L. C. Sanz. Image reconstruction from frequency-offset Fourier data. *Proceedings of the IEEE*, 72(6):661–669, June 1984.

[79] F. Natterer. *The Mathematics of Computerized Tomography*. John Wiley & Sons, New York, 1986.

[80] L. M. Novak, M. C. Burl, and W. W. Irving. Optimal polarimetric processing for enhanced target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):234–244, January 1993.

[81] L. M. Novak, G. J. Owirka, and C. M. Netishen. Performance of a high-resolution polarimetric SAR automatic target recognition system. *Lincoln Laboratory Journal*, 6(1):11–23, Spring 1993.

[82] G. J. Oxlee. *Aerospace Reconnaissance*, volume 9 of *Brassey's Air Power*. Brassey's, London, 1997.

[83] B. U. Park and J. S. Marron. Comparison of data-driven bandwidth selectors. *Journal of the American Statistical Association*, 84(409):66–72, March 1990.

[84] E. Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33(3):1065–1076, September 1962.

[85] L. C. Potter, D.-M. Chiang, R. Carrière, and M. J. Gerry. A GTD-based parametric model for radar scattering. *IEEE Transactions on Antennas and Propagation*, 43(10):1058–1068, October 1995.

[86] L. C. Potter and R. L. Moses. Attributed scattering centers in SAR ATR. *IEEE Transactions on Image Processing*, 5(1):79–91, January 1997.

[87] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2nd edition, 1992.

[88] R. Prony. Essai experimental et analytique sur les lois de la dilatabilite de fluides elastiques et sur celles de la force expansion de la vapeur de l'alkool, a differents temperatures. *Journal de L'Ecole Polytech*, 1:24–76, 1795.

[89] M. A. Rahman and K.-B. Yu. Total least squares approach for frequency estimation using linear prediction. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35(10):1440–1454, October 1987.

[90] B. D. Rao and K. Kreutz-Delgado. An affine scaling methodology for best basis selection. *IEEE Transactions on Signal Processing*, 47(1):187–200, January 1999.

[91] R. Redner and H. F. Walker. Mixture densities, maximum-likelihood estimation and the EM algorithm (review). *SIAM Review*, 26(2):195–239, April 1984.

[92] G. T. Ruck, D. E. Barrick, W. D. Stuart, and C. K. Krichbaum. *Radar Cross Section Handbook*. Plenum Press, New York, 1970. 2 volumes.

[93] J. J. Sacchini, W. M. Steedly, and R. L. Moses. Two-dimensional Prony modeling and parameter estimation. *IEEE Transactions on Signal Processing*, 41(11):3127–3137, November 1993.

[94] K. Sarabandi and T.-C. Chiu. Optimum corner reflectors for calibration of imaging radars. *IEEE Transactions on Antennas and Propagation*, 44(10):1348–1361, October 1996.

[95] J. R. Schott. *Matrix Analysis for Statistics.* John Wiley & Sons, New York, 1997.

[96] H. J. Scudder. Introduction to computer aided tomography. *Proceedings of the IEEE*, 66(6):628–637, June 1978.

[97] K. S. Shanmugan and A. M. Briepohl. *Random Signals: Detection, Estimation and Data Analysis.* John Wiley & Sons, New York, 1988.

[98] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, MI-1(2):113–122, October 1982.

[99] M. I. Skolnik. Fifty years of radar. *Proceedings of the IEEE*, 73(2):182–197, February 1985.

[100] E. C. G. Sudarshan and N. Mukunda. *Classical Dynamics: A Modern Perspective.* John Wiley & Sons, New York, 1974.

[101] C. Taswell. Satisficing search algorithms for selecting near-best bases in adaptive tree-structured wavelet transforms. *IEEE Transactions on Signal Processing*, 44(10):2423–2438, October 1996.

[102] K. Tomiyasu. Tutorial review of synthetic aperture radar (SAR) with applications to imaging of the ocean surface. *Proceedings of the IEEE*, 66(5):563–583, May 1978.

[103] F. T. Ulaby and M. C. Dobson. *Radar Scattering Statistics for Terrain.* Artech House, Norwood, MA, 1989.

[104] F. T. Ulaby and C. Elachi, editors. *Radar Polarimetry for Geoscience Applications.* Artech House, Norwood, MA, 1990.

[105] H. L. Van Trees. *Detection, Estimation, and Modulation Theory, Part I.* John Wiley & Sons, New York, 1968.

[106] J. T. Van Zandt. To live is to fly. *High, Low and In Between*, 1972.

[107] V. N. Vapnik. *Statistical Learning Theory.* John Wiley & Sons, New York, 1998.

[108] G. S. Watson. The statistics of orientation data. *Journal of Geology*, 74:786–797, 1966.

[109] D. R. Wehner. *High Resolution Radar.* Artech House, Norwood, MA, 1987.

[110] W. M. Wells III. *Statistical Object Recognition.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1993.

[111] C. A. Wiley. Synthetic aperture radars. *IEEE Transactions on Aerospace and Electronic Systems*, AES-21(3):440–443, May 1985.

[112] J. Wissinger, R. B. Washburn, N. S. Friedland, A. Nowicki, D. R. Morgan, C. Chong, and R. Fung. Search algorithms for model-based SAR ATR. In E. G. Zelnio and R. J. Douglass, editors, *Algorithms for Synthetic Aperture Imagery III*, volume 2757 of *Proceedings of SPIE*, pages 279–293, Orlando, FL, April 1996.

[113] T.-T. Wong, W.-S. Luk, and P.-A. Heng. Sampling with Hammersley and Halton points. *Journal of Graphics Tools*, 2(2):9–24, 1997.

[114] N. H. Woodcock. Specification of fabric shapes using an eigenvalue method. *Bulletin of the Geological Society of America*, 88(9):1231–1236, September 1977.

[115] N. H. Woodcock and M. A. Maylor. Randomness testing in three-dimensional orientation data. *Journal of Structural Geology*, 5(5):539–548, 1983.

[116] J. W. Woods. Two-dimensional discrete Markovian fields. *IEEE Transactions on Information Theory*, IT-18(3):232–240, March 1972.

[117] C. Wu. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11:95–103, January 1983.

[118] N. H. Youssef. Radar cross section of complex targets. *Proceedings of the IEEE*, 77(5):722–734, May 1989.

[119] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, and L. M. Novak. K-distribution and polarimetric terrain radar clutter. *Journal of Electromagnetic Waves and Applications*, 3(8):747–768, August 1989.