2012

Prediction of Optimal Bayesian Classification Performance for LADAR ATR

Kristjan H. Greenewald
Wright State University

Prediction of Optimal Bayesian Classification Performance for LADAR ATR

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Engineering

By

KRISTJAN HERBERT GREENEWALD
B.S.E.E, Wright State University, 2011

2012
Wright State University

Wright State University
SCHOOL OF GRADUATE STUDIES

August 20, 2011

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Kristjan Herbert Greenewald ENTITLED Prediction of Optimal Bayesian Classification Performance for LADAR ATR BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science in Engineering.

Brian Rigling, Ph.D.
Thesis Director

Kefu Xue, Ph.D.
Department Chair of Electrical Engineering

Committee on
Final Examination

Brian Rigling , Ph.D.

Bill Pierson , Ph.D.

Fred Garber , Ph.D.

Andrew T. Hsu, Ph.D.
Dean, Graduate School

ABSTRACT

Greenewald, Kristjan H., M.S.Egr, Department of Electrical Engineering, Wright State University, 2012 . *Prediction of Optimal Bayesian Classification Performance for LADAR ATR.*

We consider the problem of LADAR ATR classifier performance prediction in the presence of arbitrary nuisance parameters including but not limited to pose. We use several noise models for both range images and point clouds that are significantly more accurate and complex than the Gaussian models used by previous non-Monte Carlo prediction methods. Two accurate new methods of efficiently predicting the optimum Bayesian classification performance are then derived, and applied to the noise models. Advantages of these methods include significant gains in accuracy for medium to high noise levels and the ability to handle target near symmetry. Extensions are developed for multiple targets and predicting the performance of classifiers designed using incorrect noise models. We also derive several simple analytic approximations for the behavior of the probability of error as important sensor and noise parameters vary. Finally, we verify the accuracy of our predictions using Monte Carlo simulations.

Contents

1 Introduction **1**

1.1 Motivation . . . 2

1.2 Contributions . . . 3

1.3 Thesis Outline . . . 4

2 Background **5**

2.1 Automatic Target Recognition . . . 5

2.2 Target Classification . . . 6

2.3 ATR Performance Prediction . . . 7

2.4 LADAR ATR Performance Prediction . . . 8

2.5 LADAR . . . 10

2.5.1 Sensor Description . . . 10

2.5.2 Types of LADAR . . . 11

2.5.3 Output Image Types . . . 12

2.6 LADAR Noise . . . 13

2.6.1 Noise Models for Performance Prediction . . . 13

2.6.2 Noise Models for Image Synthesis . . . 14

3 LADAR Noise Models **16**

3.1 LADAR Image Synthesizer . . . 17

3.2 Noise Models . . . 17

3.2.1 Gaussian Noise with Anomalies - Coherent LADAR . . . 18

3.2.2 Gaussian Noise with Anomalies- Geiger mode LADAR . . . 19

3.2.3 Gaussian Noise with Random Deletions . . . 23
3.2.4 Finite beamwidth effects . . . 24

4 Bayesian Classification Performance Prediction 27

4.1 Introduction . . . 27
4.2 General Bayesian Hypothesis Testing under Nuisance Parameters . . . 28
4.3 Method of Grenander et. al. . . . 29
4.3.1 Approximation of Bayesian Integral using Laplace's method . . . 32
4.4 Method of Noiseless Integration . . . 33
4.5 Unbiased Integration Method . . . 35
4.6 Prediction of Probability of Error . . . 40
4.7 Extensions . . . 40
4.7.1 Multi-modal Extension . . . 41
4.7.2 Extension to M-ary Classification . . . 43
4.7.3 Extension to Incorrect Classifier Noise Models . . . 44

5 Implementation 46

5.1 Computing the Integrals . . . 46
5.1.1 Method of Grenander . . . 46
5.1.2 NIM and UIM . . . 47
5.1.3 Sample Selection for Integration . . . 48
5.2 Computing the Pseudo LLR PDF . . . 49
5.3 Approximating the Accuracy of the Prediction . . . 51

6 Application to LADAR Noise Models 53

6.1 Noise Model Components . . . 54
6.1.1 Gaussian Noise . . . 55
6.1.2 Mixture PDF . . . 57
6.1.3 Random Deletions . . . 59
6.2 LADAR Noise Models . . . 64
6.2.1 Gaussian Noise with False Alarms - Coherent LADAR . . . 64

6.2.2 Gaussian Noise with Geiger mode anomalies . . . 65
6.2.3 Gaussian Noise with Random Deletions . . . 67
6.2.4 Finite beamwidth effects . . . 68

7 Analytic Parameter Variation Approximations 70

7.1 Asymptotic Resolution Dependence for Independent Pixel Noise . . . 70
7.2 Analytic Approximations for Gaussian Noise with Anomalies . . . 75
7.2.1 Probability of Anomaly Dependence . . . 75
7.2.2 Noise Covariance Dependence . . . 76
7.2.3 Elevation Angle Dependence . . . 77

8 Bayesian Experimental Results for LADAR Noise 80

8.1 1-D Gaussian Noise . . . 80
8.2 1-D Gaussian with Anomalies- Coherent . . . 82
8.3 1-D Gaussian Noise with Geiger-mode Anomalies . . . 83
8.4 1-D Gaussian Noise with Random Deletions . . . 84
8.5 3-D Gaussian Noise with Random Deletions . . . 85
8.6 Coherent LADAR with finite beamwidth effects . . . 87

9 Experimental Studies 92

9.1 Performance for Multiple Target Classification . . . 92
9.2 Performance for Multiple Nuisance Parameters . . . 95
9.3 Target Symmetry . . . 96
9.4 Multi-Modal Extension . . . 98
9.5 Asymptotic Dependence of Performance on Resolution . . . 100
9.6 Dependence of Performance on Anomaly Probability . . . 103
9.7 Performance as a Function of Sensor Elevation Angle . . . 106
9.8 Performance as a function of Angular Pose . . . 110
9.9 Pixel Contributions to Performance . . . 113

10 Conclusion 120

A Proof of Theorem I **123**

B Asymptoticity of Deletion Approximation for Coherent LADAR Anomalies **125**

Bibliography **128**

List of Figures

3.1 Noiseless Range Image 18

3.2 Coherent LADAR Noise 20

3.3 Geiger mode LADAR noise 23

3.4 Linear mode LADAR noise with finite beamwidth 26

4.1 View of Tank CAD model from simulated LADAR sensor location. 34

4.2 View of Sedan CAD model from simulated LADAR sensor location. 35

4.3 Plot of integrand for low noise 36

4.4 Plot of integrand for high noise 37

8.1 Probability of error for 1-D Gaussian noise 82

8.2 Probability of error for coherent LADAR model 83

8.3 Probability of error for Geiger LADAR model 85

8.4 Probability of error for Gaussian noise plus random deletions 86

8.5 Probability of error for 3-D Gaussian noise with random deletions 87

8.6 Probability of error for finite beamwidth 89

8.7 Probability of error for zero beamwidth 90

8.8 Probability of error for incorrect noise model 91

9.1 Honda Civic Model, shown in a noiseless LADAR range image. 93

9.2 Toyota Avalon Model, shown in a noiseless LADAR range image. 94

9.3 Mazda Sentra Model, shown in a noiseless LADAR range image. 95

9.4 Jeep Model, shown in a noiseless LADAR range image. 96

9.5 Probability of selection of Avalon 97

9.6 Probability of selection of Civic . . . 98
9.7 Probability of selection of Sentra . . . 99
9.8 Probability of selection of Jeep . . . 100
9.9 Performance for Three Nuisance Parameters . . . 101
9.10 Probability of error for full width prior and uniform anomalies . . . 102
9.11 Full width prior approximated with $W = 2$. . . . 103
9.12 Full width prior approximated with $W = 4$. . . . 104
9.13 Probability of error as function of resolution . . . 105
9.14 Probability of error as function of resolution: simple asymptotic approximation . . . 106
9.15 Probability of error as function of resolution: better asymptotic approximation . . . 107
9.16 Probability of error as function of probability of anomaly: Random Deletions . . . 108
9.17 Approximation for performance variation with anomaly rate: Coherent LADAR . . . 109
9.18 Performance as a function of elevation angle . . . 110
9.19 Approximate performance as a function of elevation angle . . . 111
9.20 Approximate performance as a function of elevation angle . . . 112
9.21 Probability of error as function of truth azimuth angle . . . 113
9.22 View of true target (Civic) in the pose for which the performance is best . . . 114
9.23 View of true target (Civic) in the pose for which the performance is worst . . . 115
9.24 Number of pixels on target as a function of pose . . . 116
9.25 Normalized probability of error as function of truth azimuth angle . . . 117
9.26 Normalized expected LLR as function of truth azimuth angle . . . 118
9.27 View of Civic in pose used for individual pixel analysis. . . . 118
9.28 Expected value of pseudo loglikelihood ratio . . . 119
9.29 Pixel contributions to the Laplace approximation threshold adjustment. . . . 119

Acknowledgement

I would like to take this opportunity to extend my thanks to my advisor, Dr. Brian Rigling, for his guidance and insights throughout my time at Wright State University. In addition, I appreciate his patience in critiquing and helping me to revise my work. Without all those revisions and feedback, this work could not have been what it is. I would also like to thank Drs. Fred Garber and Bill Pierson for taking the time to serve on my committee, and for providing excellent insights. I would like to thank my mentors at the Air Force Research Laboratory (AFRL) Sensors Directorate, especially Ed Zelnio for his vision and guidance for this work, and Linda Moore for her help throughout. Finally, I would like my family for their support and encouragement, without which I would not have been able to complete this work.

Dedicated to

My Parents

Chapter 1

Introduction

ATR performance prediction is an essential goal in the development of ATR theory. For many applications, it is of particular interest to be able to approximate the performance of the optimum classifier, in order to provide a computationally inexpensive method of evaluating the relative quality of proposed ATR algorithms, or to provide direction how to adjust system parameters in order to improve an ATR system. Other applications include sensor management, for which it is important that the ATR system is able to know how ATR performance will be affected as parameters change, so as to better select those parameters. For this thesis, the ATR performance metric of interest is the probability of correct classification, which, unfortunately, is relatively difficult to compute exactly in realistic situations. This is primarily due to the presence of unknown nuisance parameters such as target pose that have significant effects on the appearance of the target. This aspect of the problem also makes performance evaluation using Monte Carlo simulations particularly expensive in most cases. Thus, the primary challenge of ATR performance prediction lies in incorporating these nuisance parameters into the prediction in an accurate and computationally efficient manner.

A variety of imaging sensors are used for ATR. With improved FLASH LADAR technology, LADAR sensors are increasingly being used to perform ATR, especially when other sensors such as radar are unable to definitively identify the target. LADAR sensors provide high-resolution, albeit noisy, three-dimensional images of the side of the target visible from the LADAR sensor, thus making it a desirable choice for ATR tasks that involve distinguishing similar targets.

1.1 Motivation

In this thesis, we focus on the prediction of the classification performance of LADAR ATR. In general, the ATR problem consists of identifying the class of a target in a given image. A list of possible target classes $\mathcal{L} = \{\alpha_1, ..., \alpha_M\}$ are known to the algorithm, and hypothesis testing is used to determine which class the target belongs to, or if it is from an unknown class. The distribution of the noisy image given a particular target typically varies as a function of a set of nuisance parameters that are unknown to the classifier. Since the exact pose and position of the observed target have a significant impact on the observed image and are not available to the classifier, the pose and position of the target are important nuisance parameters in virtually all ATR scenarios. Other common nuisance parameters are related to target configuration.

Previous work in LADAR ATR performance prediction under nuisance parameters has made use of the Laplace approximation [1, 2], which involves a simple modification of the threshold on the likelihood ratio test to correct for uncertainty in different nuisance parameter estimates for different targets. This approximation holds for the asymptotic case where noise levels decrease to zero. To make use of this approximation, techniques have been developed using numerical techniques to compute the Cramer-Rao bounds and similar measures for the pose parameters involved in LADAR range images and other imagery [3–7]. For similar imaging sensors, performance prediction has also been done using information theoretic quantities [8] to bound the performance, although these bounds tend to not be tight. Moreover, these methods so far have focused on overly simplistic noise models, which ignore anomalous pixels and cross range blurring and noise, thus overestimating the probability of correct classification. They also make inaccurate assumptions, such as that no rotational near symmetries are present in the target. In addition, these methods focus on range images only, in spite of the fact that point clouds are increasingly becoming the LADAR image format of choice. As a result, Monte Carlo methods are typically used to evaluate performance. These methods are time consuming and thus make it difficult to get accurate performance predictions.

1.2 Contributions

We focus on obtaining more realistic and accurate predictions of LADAR performance. First, we review the LADAR noise modeling and image synthesis literature and develop more accurate LADAR noise models for both range images and point clouds, while ensuring that they remain simple enough to obtain noise pdfs, along with some simplified models that allow for closed form solutions. Proposed noise models with increased accuracy include a Geiger mode anomaly model, a point cloud noise model, and finite receiver beamwidth and beam divergence noise models. To make it easier for analysis to be performed while retaining anomalous pixels, a model of anomalous pixels as random deletions is also used. These noise models were implemented in a modified LADAR imagery synthesis software package.

Secondly, we derive improved methods of predicting performance. We consider two new methods of predicting the performance of the optimal Bayes classifier with arbitrary pixel noise. Since the Bayesian classifier is theoretically optimal, the predictions are an approximation to the upper bound of performance for all algorithms for the same problem. The first method is a simple extension of the method in [1] for arbitrary noise and increased accuracy. The second method is a new method that selects the threshold adjustment in a way that involves an unbiased approximation of the Bayesian integral, thus correcting the bias of the first method. Both methods significantly outperform the Laplace approximation methods, with the second method outperforming the first, frequently by a large margin, especially in high noise situations. The computational complexity of the prediction remains quite low, and is roughly the same as that of performing one Bayesian classification of an observed target. A multiple target extension of these methods is then developed, as well as important extensions for the computation of the performance of Bayesian classifiers derived using incorrect noise models. Finally, a method of improving the accuracy of the predictions is presented, where, by gradually increasing the prediction complexity, greater accuracy can be achieved. Both of these proposed prediction methods are then applied to each of the noise models considered.

Simple analytic and near analytic techniques of approximating the variation of the probability of error with respect to the parameters of sensor resolution, probability of anomaly, sensor elevation angle, and noise variance are then considered. As opposed to previous results [2], many of these

techniques are based on extrapolating from the predicted performance at one value of the parameter to the performance at another value. This allows an accurate performance prediction method to be used to predict the performance at one or more points on the performance curve, and then these simpler approximations can be used to efficiently fill in the rest of the curve. As a result, these methods may be more useful in that by extrapolating from a known point on the curve, the error due to the simplifying assumptions is mitigated as opposed to attempting to compute the entire curve using the assumptions with no reference to more accurate prediction methods.

1.3 Thesis Outline

The remainder of the thesis is organized as follows. In Chapter 2, we review the relevant literature and discuss LADAR ATR and the LADAR sensor and models for it. In Chapter 3, we develop and present the noise models used in this thesis. In Chapter 4, we discuss the existing Laplace approximation Bayesian prediction method, and present our two prediction methods and their extensions. In Chapter 5, we consider methods of efficiently implementing the various parts of our prediction methods. Chapter 6 derives the application of these new methods to the noise models presented earlier. In Chapter 7, approximate analytic results for the performance variation due to common parameter variations are considered. Chapter 8 presents some experimental results for the Bayesian prediction methods and compares them to Monte Carlo simulation results to verify the accuracy of the predictions. Finally, in Chapter 9, we present some additional prediction results, focusing on the variation of performance with various parameters and the influence of nuisance parameters and statistical difference to LADAR ATR performance. We present our conclusions in Chapter 10.

In this thesis, we use the following notation. Scalars are denoted by italic letters, such as a. Vectors are denoted by underlined variables, such as $\underline{a}$, and matrices by boldface capital letters, such as $\mathbf{A}$. Sets are denoted by script capitals, such as $\mathcal{A}$. Matrix inverse is denoted by $\mathbf{A}^{-1}$, and transpose is denoted by $\mathbf{A}^{T}$.

Chapter 2

Background

In this chapter, the background information and literature review necessary for the development of this thesis is presented. In Section 2.1, the automatic target recognition (ATR) problem is introduced. A brief discussion of typical ATR classification methods is presented in Section 2.2. ATR performance prediction is introduced in Section 2.3, along with some relevant performance prediction methods used for non-LADAR ATR. Section 2.4 presents a review of existing LADAR ATR performance prediction methods. Next, the LADAR sensor is introduced and described in Section 2.5, and a survey of LADAR noise models is presented in Section 2.6.

2.1 Automatic Target Recognition

There are many challenges in ATR. Since the targets of interest are present in a frequently complex scene, the targets must be detected and segmented from the background, and ATR classification must thus be robust to segmentation errors and obscuration. Due to the enormous variation of background scenes, as well as the large number of non-targets and confusers that are somewhat similar to targets of interest, the classifier must also determine if the observed target is not a member of the set of targets it can classify.

The presence of nuisance parameters is a major issue with ATR, since from a classification standpoint the distribution of each class varies dramatically with many nuisance parameters, espe-

cially for pose [1, 2]. Nuisance parameters are parameters that affect the distribution of the image of each class. Common nuisance parameters are target position and orientation. Other nuisance parameters include configuration, or within class variability. Because of these nuisance parameters, basic likelihood ratio classifiers are typically not directly applicable. As a result, other methods of classification have been developed.

2.2 Target Classification

The theoretically optimal method for classification of targets in the presence of nuisance parameters is the Bayesian maximum a posteriori (MAP) classifier [9]. It, however, has historically been too complex to implement in practice due to the necessity of integrating out the nuisance parameters from the observed image likelihood functions for each class. This integral must frequently be calculated numerically, using a noise model and a model of the target appearance. As a result, several approximations to the Bayesian integral have been employed, such as the Laplace approximation [1] to simplify the calculation.

In practice, however, the Bayesian approach is frequently not the most straightforward. Noise models are frequently overly simplistic [1] and priors on the nuisance parameters must be selected. Further complicating the situation is the fact that the background is completely unknown, so Bayesian methods cannot take it into account.

As a result, other techniques of classification have been developed. One of the earliest methods applied to ATR was the generalized likelihood ratio test (GLRT) which involves substituting the MLE of the nuisance parameter vector into the likelihood for each hypothesis and computing their ratios to the threshold. Invariance and other feature-based techniques are often used as well [10–12]. The invariance approach is feature based, in that it selects several functions of the detected target, and uses these values to classify it instead of the entire image [10]. To avoid having to work with nuisance parameters, invariant features are selected, that is, they are chosen such that they ideally do not vary with nuisance parameter value. Regardless of the choice of features, since the distributions of the feature vectors is usually difficult to derive analytically, the distributions

are typically estimated using an appropriate dataset and a classifier designed. When models of the targets are available, features based on "distances" to these models can also be used [11, 12].

2.3 ATR Performance Prediction

In this section, we examine previous work with ATR performance prediction. ATR performance prediction is an important aspect of ATR theory, in that it enables ATR systems to be designed more efficiently using the insights and estimates provided, as well as providing a benchmark with which to compare ATR algorithm performance.

Classifier performance is typically measured by the probability that the classifier commits an error [13]. This includes the conditional error probability, which is the probability that an error is made given that the true target is target α with a particular set of nuisance parameters, and the unconditional error probability, which is the overall probability that an error will occur. Methods of predicting classifier performance reviewed here include information-theoretic bounds, Laplace approximation methods, and individual algorithm predictions.

In [1], Grenander, Srivastava, and Miller consider the problem of classifier performance prediction for ATR. Since the Bayesian classifier is optimal in the sense that it minimizes the resulting probability of error, they predict the performance of the Bayes MAP classifier, with the nuisance parameters removed from the hypothesis likelihood functions by integration. The integration of the nuisance parameters for the Bayesian classifier creates a likelihood ratio that is the ratio of two random integrals. Since all parts of the integrand are correlated, determining the exact pdf of the integrals is difficult to do analytically in most cases. The authors thus seek an approximation. The authors propose a prediction method that they demonstrate is asymptotically correct as the noise variance approaches zero. For the binary classification scenario, it is based on creating a simple likelihood ratio between the true target with the true values of the nuisance parameters and the incorrect target with the nuisance parameters that maximize the likelihood of the true target given the incorrect target.We discuss the derivation of this prediction method in a later chapter based on approximating the Bayesian integral using the Laplace low-noise approximation.

In [8], Jain, Moulin, Miller, and Ramchandran apply information-theoretic bounds on probability of error to the problem of optical image ATR. The system model used by the authors is that the image is composed of a target corrupted by clutter and pixel noise. To obtain the bounds on the probability of error, the authors turn to the standard information-theoretic distance measures in the Ali-Silvey class [8]. They focus primarily on the Chernoff distance, which provides an upper bound on both conditional probabilities of error. Two methods for obtaining a upper bound on the probability of error for detection under nuisance parameters are presented. It is found by experiment that the upper bounds do indeed hold, but are not very tight, especially for one CAD target considered. A very simple asymptotic prediction is also shown, and it greatly outperforms the Chernoff results in terms of accuracy. Finally, Jain et al. extend these results to multi-target ATR. This approximation is not valid for the case where more than two of the targets are relatively close together. The authors then use the Chernoff bound to bound the probabilities of error, and derive an upper bound for the Chernoff distance in the form of a double integral with respect to the nuisance parameter space.

In [14], Garber and Zelnio consider ATR performance prediction for radar. Their method is based on communication theory techniques relating to sensor and target "capacity" to obtain a rough approximation to the probability of error. Their results are derived for the binary template matching of scattering centers in radar range profiles of the target. Their results achieve a rough estimate of performance, as the predicted curves are generally within an order of magnitude of the truth.

In [15], Vore considers SAR ATR performance for the case where the parameters of the classifier's model-based target distributions are incorrect. The author makes the simplifying assumption that the image distributions can be viewed as observation vectors that are complex Gaussian, and then derives the probability of error when the classifier is based on incorrect target means and covariances. In all cases, the predicted error rates are off by at least a factor of two.

2.4 LADAR ATR Performance Prediction

Due to its high resolution and its provision of 3-D information, LADAR provides a rich dataset for target recognition. Many algorithms have been proposed and studies performed for LADAR ATR

classification. In this section, we focus on target classifier performance and nuisance parameter estimation performance.

Yen and Shapiro in [2] derive asymptotic expressions for LADAR ATR performance prediction. Their method is based on the Laplace approximation method used in [1] to predict the performance of the optimal Bayesian classifier. The recognition problem considered is the identification of targets with random pose with a uniform prior. Their LADAR model assumes IID pixel noise and composed of a mixture of a Gaussian pdf and a uniform pdf, where the uniform pdf models the anomalies in the imagery. To simplify their analysis, however, the authors assume that the range interval is so short that the probability of anomaly can be approximated as zero, thus effectively returning to the AWGN noise model of [1].

The authors then use the AWGN Laplace approximation in [1] to obtain the conditional probability of error, that is, the probability of error given that the true target has a particular set of nuisance parameters unknown to the classifier.

To obtain the unconditional probability of error, that is, the probability of error given that the nuisance parameters take on random values according to a uniform prior, the authors use the exponential approximation of the Q function and the unconditional error results. The authors also derive an expression for the variation of the probability of error with changes in sensor resolution for FLIR. They assume that the pixel observations remain independent, and the resolution is high enough that nearby pixels have virtually the same noiseless pixel value, which is either one or zero. Using this and the asymptotic unconditional error probability, they obtain an approximate analytic result.

Koksal and Shapiro in [7] consider analytic approximations to the Hilbert-Schmidt bound on orientation estimation using both LADAR and FLIR. Since orientation angle is not a flat Euclidean space, the Cramer Rao bound on estimation performance is not valid. Instead, a Lie group representation should be used to capture the inherent periodicity. The appropriate MMSE estimator can then be created by minimizing the Hilbert-Schmidt norm associated with the parameter space. The mean squared error achieved by this estimator is the minimum possible, and the authors refer to it as the Hilbert-Schmidt bound. They consider a specific blocks-world target (e.g. composed of rectangular

blocks only, all in the same orientation). Using the standard coherent LADAR model with the assumption that all anomalies can be perfectly identified, they were able to analytically calculate the Cramer Rao bound on the orientation estimation for the target, and using this, compute an analytical approximation to the Hilbert-Schmidt bound for LADAR, up to a constant which depends on the target geometry. All the analytical results were compared with Monte Carlo simulation results and found to be reasonably accurate.

In [3], Dixon and Lanterman consider the numerical calculation of Cramer Rao bounds on ground target pose estimation for laser radar. They use the same Gaussian and uniform mixture noise model as proposed by [2]. Since anomalies do not provide any information about the target, they ignore them in their derivation. To the nuisance parameter of angular orientation considered by Koksal and Shapiro, they add the nuisance parameters of (x, y) ground position as the parameters to be estimated. For the Gaussian noise model, they derived a method of using synthetically generated LADAR imagery, log-likelihood function evaluation, and computation of derivatives using finite differencing to obtain the Fisher information matrix. The authors then used this method to numerically compute the CRLBs for CAD model targets, as a function of range and orientation angle. It was found that the CRLB decays with target range from the sensor, and varies significantly with orientation angle. The authors then use this observation to argue that invariant ATR methods must therefore be suboptimal.

2.5 LADAR

2.5.1 Sensor Description

LADAR sensors create pixel-based images of stationary scenes [16]. The pixel values are the measured range along an angle-angle line of sight from the sensor [16]. The range information is obtained by emitting a short pulse from the sensor to the scene, detecting the reflection, and measuring the resulting time of flight, which is then used to obtain a range measurement based on the speed of light [16]. Except for synthetic aperture LADAR [17], which we do not consider, cross-range

resolution is obtained by limiting the field of view of the individual receivers or the laser beam, thus limiting the area of the scene whose reflections can reach the sensor [16].

There are several types of LADAR sensors. In this thesis, we focus on the use of range images and point clouds only, as is commonly done with LADAR performance prediction [2, 3]. These observations are corrupted by noise due to a variety of factors such as laser speckle, sensor timing noise, shot noise, and atmospheric turbulence [17–20]. One of the characteristic features of LADAR noise is the presence of dropouts and anomalies [19–21]. Since the images are formed by detecting reflected light from the scene, it is quite common for an individual pixel range measurement to be lost because the sensor failed to detect the reflected light, or to become an anomaly because background light was detected instead of the true reflected light, resulting in the measurement having nothing to do with the true range [19–22]. To reduce anomalies and limit the extent of the observed scene, LADAR sensors also employ range gates, thus blocking all detections that are outside a specified interval of range values that constitute the "range ambiguity interval" [19–22].

2.5.2 Types of LADAR

LADARs generally use one of two methods to obtain range measurements [16, 22]. The first type, which is also the oldest, is called scanning LADAR [16]. In this method, the laser beam sequentially scans over the region of interest, sending out pulses at specified intervals, usually creating a rectangular raster pattern. For each pulse, the reflected laser light is detected using a single receiver. Since the beam is scanned, very high cross-range resolution can be achieved.

Most modern LADARs are flash LADARS. Flash LADARs achieve cross-range resolution using an array of receivers closely spaced in angle [22]. Instead of scanning the laser beam across the scene, the laser pulse illuminates the entire scene. When the reflected pulse returns, each receiver detects the portion of the laser beam that reflected off the small portion of the scene at which it was pointed [17, 22]. The area of this region is limited by the beamwidth of the receiver.As this method allows multiple pixel observations to be measured at the same instant, the speed at which images can be produced is considerably higher than that of scanning LADAR [22, 23]. This comes at the cost of cross-range resolution limited by the APD spacing on the array.

The array of APDs on a flash LADAR is usually relatively small (e.g. 16×16 or 64×64) [24], and insufficient for many image processing tasks. As a result, flash LADARS are frequently used to take a large number of successive images of the scene over a short period of time. These small images are then combined into a large image with a much higher number of pixels.

Geiger-mode direct-detection LADARs use Geiger-mode photon-counting avalanche photodiodes (APDs) to detect individual photons reflected from the scene [20, 23, 25, 26]. A primary benefit of using these APDs is that the light detection sensitivity, and thus the detection probability, of the LADAR is significantly higher than it would be for coherent-detection LADAR or older direct detection LADAR [19]. This allows the LADAR to be used at greater ranges and lower laser powers, but on the other hand, results in an increased likelihood that background light will be detected, thus creating anomalous pixels. Furthermore, Geiger-mode LADARs are typically used as a flash LADAR instead of as a scanning LADAR, which is the primary imaging method used for coherent LADARs.

Geiger-mode APDs have the characteristic property that after they are activated, they fire as soon as the first photon impacts its surface, and then cannot detect another photon for some time [20, 23]. This is opposed to the coherent LADAR detectors, which record the reflected light intensity for the entire time interval, and select the highest intensity point for its range measurement [21]. As might be expected, this can cause Geiger-mode LADAR to have significantly higher anomaly rates, thus usually requiring multiple images taken in succession to properly resolve the scene [22]. Secondly, this results in anomalies being biased towards the region of the range interval closest to the sensor [20, 23].

2.5.3 Output Image Types

LADAR sensors measure both range and intensity information [17]. The range information is typically output either as a 2-D range image, for which each pixel corresponds to a different line of sight emanating from the sensor, and whose value is equal to the measured range to the scene, or as a 3-D point cloud [1, 2, 22]. The lines of sight are typically arranged so that the angular spacing between pixels is such that a uniform grid is formed [1, 2].

The 3-D point cloud image presents the same data presented in a range image, but for each pixel computes the (x, y, z) position of the observed point in the scene, computed from the pixel's range value and the known line of sight associated with the pixel [18]. The added benefit of using this type of image is that it is easier to fuse a series of LADAR images together to form one 3-D image, as is frequently done in modern LADAR image formation, especially for flash LADAR [18, 25].

2.6 LADAR Noise

2.6.1 Noise Models for Performance Prediction

In the LADAR performance prediction literature, it is common to use noise models associated with coherent scanning LADAR [2, 3, 7, 21, 27, 28]. A range image noise model that models the image as composed of independently distributed pixels is typically used. Some of these are anomalies and some of these are noisy measurements of the true associated range. The occurrence of an anomaly is modeled as being Bernoulli distributed with a specified probability of anomaly α for each pixel determined by the sensor parameters, the range, and the atmospheric state [3]. Anomalous pixels are caused by false detections occurring due to background radiation or detector noise, and thus have no relation to the true distance to the scene. This model models anomalous pixels as being uniformly distributed between the limits a and b of the sensor range gate. A non-anomalous pixel is modeled as being Gaussian distributed with mean equal to the true range value and a given (constant) variance determined by the sensor parameters [3, 27].

This gives an overall distribution of the kth pixel s_k of the image to be

$$p\left(s_k\right) = \alpha u(s_k) + \frac{1-\alpha}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(s_k - \overline{s}_k)^2}{2\sigma^2}\right\} \tag{2.1}$$

where

$$u(t) = \begin{cases} \frac{1}{b-a} & a < t < b \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

and $\overline{s}_k$ is the mean of the kth pixel and $[a, b]$ is the range interval set by the sensor, and σ is the standard deviation of the Gaussian component.

This formulation assumes that $\sigma \ll (b - a)$, $\overline{s}_k - a \gg \sigma_k$, and $b - \overline{s}_k \gg \sigma$ such that the integral of the Gaussian portion of the pdf outside of the interval $[a, b]$ is essentially zero.

Although this noise model is a reasonable one for coherent LADAR, advances in LADAR technology have resulted in the development of photon-counting LADAR imagers, which tend to produce point clouds rather than range images [22, 29]. The model is somewhat simplistic as well. For example, it assumes that cross-range noise is negligible. This is not accurate [17], however, for high pixel densities and results in image edges being sharper than in reality [17]. This results in an overestimation of the ability to estimate the pose of the target.

Some authors further simplify this noise model for performance prediction by setting the probability of anomaly to zero [2,3], or, for Hilbert Schmidt bound analysis, by approximating anomalies as random deletions [7].

2.6.2 Noise Models for Image Synthesis

In [21, 23], the authors consider the detection statistics for Geiger mode LADAR. Their analysis is based on the sensor parameters and the assumption that the range interval is short enough that once a photon is detected by a Geiger-mode APD, the APD cannot detect a second photon until the next image. The detection statistics are not the only source of noise. The local oscillator of the sensor and the behavior of the APD also introduce Gaussian noise.

Atmospheric turbulence is also a factor in LADAR noise, especially at long range. As the laser beam and its reflection travel through the air, turbulence diverts and spreads the beam slightly, and slightly attenuates it as well [17, 18, 30]. Complex simulation methods using a series of Gaussian phase screens have been developed to model this [17, 18]. Others have claimed that for most situations it is sufficient to model turbulence effects as a reduction in the probability of detection and as causing beam broadening [17, 30]. This last approximation is based on the fact that the range return is a detection from a particular point illuminated by the beam, with the exact point randomly

selected according to the average returned power. As a result, random deflection of the beam is essentially equivalent to a wider beam on average [17].

The fact that each receiver has a finite beamwidth introduces uncertainty as to the point of the target from which the detected return photon is received [17, 18]. The probability distribution function of the points in the field of view that may be the returning point is determined by the average returned power density from each point, which is in turn determined by the receiver attenuation associated with each angular portion of the beam, the angle of arrival, atmospheric attenuation, and the reflectance of the scene [17, 18]. In the case of Geiger mode noise, the nature of the detector affects this as well since it results in closer portions of the target being more likely detected.

One proposed method of discretely simulating this effect is to divide each receiver beamwidth into a number of sections [17, 18]. A certain proportion of the returned power response is associated with each section. The range associated with each section is then calculated, as well as any other relevant parameters, and the probability of each section being chosen is determined [17, 18]. This is then converted into a range pdf, to which is added any other noise such as anomalies and Gaussian range noise. As this is fairly simple to implement as a mixture distribution, this is the basis of our finite beamwidth noise models.

Chapter 3

LADAR Noise Models

In this chapter, we introduce and develop the LADAR noise models that our performance prediction methods will be applied to in this thesis. First, we introduce the concept of using a random pixel deletion model as an approximate model of anomalous pixels. We then present several noise models of varying accuracy and complexity for both coherent and Geiger-mode LADAR range images. Finally, we discuss noise models for 3-D LADAR point clouds. The actual noise model used in a particular problem will be a function of the type of LADAR imager under consideration and the desired tradeoff between accuracy, ease, and computational speed of the implementation in a performance prediction setting.

The general imaging model is as follows. Consider the case where a LADAR image $\mathcal{I}$ is a set of N pixels. For a range-image LADAR model, where the value of each pixel is the appropriate measured range to the scene, $\mathcal{I} = \{s_1, ..., s_N\}$. For a point cloud model, where the value of each pixel is the (x, y, z) coordinates of a measured scene point, $\mathcal{I} = \{\underline{s}_1, ..., \underline{s}_N\}$, where each $\underline{s}_k$ is a (3×1) vector. Let the noiseless images $\overline{\mathcal{I}}_i(\underline{\theta})$ corresponding to target hypothesis H_i with nuisance parameters $\underline{\theta}$ have pixels $\overline{\underline{s}}_{k,i}(\underline{\theta}),\ k = 1, ..., N$. Thus, assuming hypothesis H_i and nuisance parameters $\underline{\theta}$, the pixel observations $\underline{s}_k$ are distributed according to some distribution

$$p(\mathcal{I}|\overline{\mathcal{I}}_i(\underline{\theta}), H_i, \underline{\theta}) \tag{3.1}$$

A common feature of these approximate models is the assumption that the noise associated with each pixel is independent of that of the other pixels. This assumption is fairly standard in LADAR performance prediction [1–3, 7, 31] and ensures that the attractive mathematical properties of independence, such as pdf separability and simple computation of the pdf of added variables of known distributions, can be employed to reduce the computational complexity of associated operations.

3.1 LADAR Image Synthesizer

For generating all noiseless images, we use a modified version of the GTRI LADAR Simulator software [32]. This software generates noiseless LADAR images using efficient CAD model 3-D projection computation techniques. The open-source software runs in a MATLAB application, and can generate synthetic LADAR images of targets and scenes in arbitrary poses and from arbitrary viewing parameters. For the generation of noiseless images, our chief modification of the software was the addition of the capability of generating long sequences of images with the pose parameters slowly incremented. This capability is crucial for the computation of the performance predictions discussed in this thesis.

The software was also modified so as to be able to generate noisy LADAR images using a variety of noise models. Sensor imaging models implemented include Gaussian blurring, atmospheric attenuation, additive Gaussian noise, anomalous pixel noise for both Geiger and linear mode LADAR, random finite beamwidth effects via Gaussian mixture distributions, and any combination of the above.

A sample noiseless range image of an urban scene viewed from the air is shown in Figure 3.1.

3.2 Noise Models

In this section, we discuss noise models for LADAR. As Gaussian range jitter is always present in LADAR, each of the noise models presented here is based on a mixture of a Gaussian pdf with some

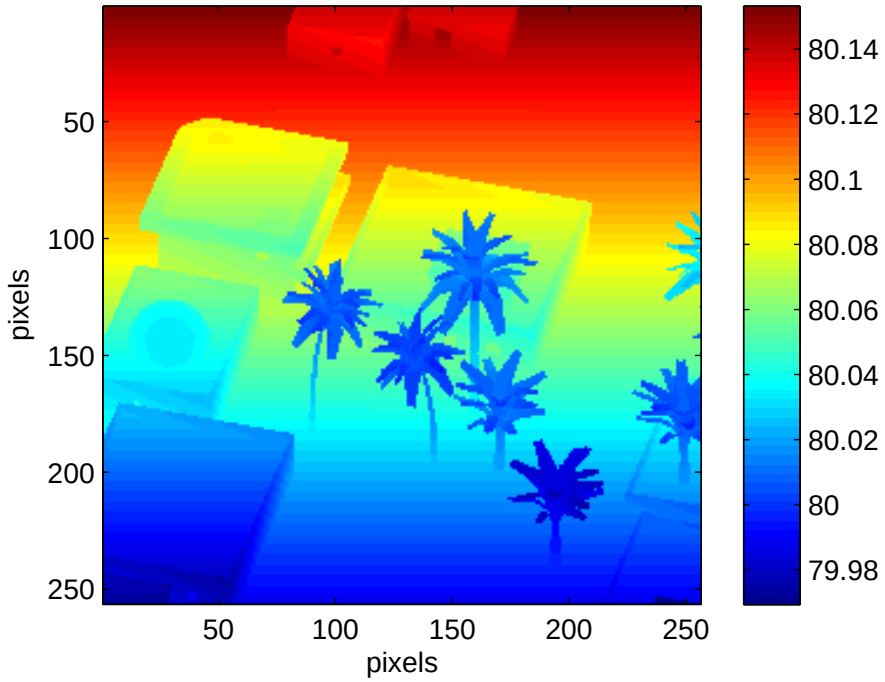

Figure 3.1: Noiseless 256×256 pixel range image of urban scene. Longer ranges are red and shorter ranges are blue.

anomaly distribution model. We discuss the typical coherent LADAR noise model, a Geiger-mode LADAR model, and a simpler random deletions model for anomalous pixels and use this to create overall noise models. Finally, we introduce a model for cross range noise applicable to each of the anomalous pixel models.

3.2.1 Gaussian Noise with Anomalies - Coherent LADAR

For coherent-detection LADAR and range images, we use the noise distribution proposed by Shapiro and Green [27, 28], and used by others in the literature [2, 3, 7, 21]. In this noise model, each range image pixel is distributed according to a pdf that is a mixture of a uniform pdf and a Gaussian. This model is also accurate for linear-mode direct detection LADARs [21, 23]. The uniform pdf models the anomalous pixels, which are caused by random laser speckle [2, 21], while the Gaussian pdf models the pixels that are observing the scene. The Gaussian noise is caused by random local oscillator shot noise [2, 21]. It is assumed that the noise distribution for each pixel is independent

of that of every other pixel. This assumption is valid for the case where the LADAR pixel spacing is large enough that the laser speckle is virtually uncorrelated from pixel to pixel. Following the general scenario where all noise parameters can vary from pixel to pixel, the pdf of the kth pixel under hypothesis H_i is given by

$$p\left(s_k|H_i\right) = \begin{cases} \frac{\alpha_{k,i}}{b-a} + \frac{1-\alpha_{k,i}}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(s_k-\overline{s}_{k,i}(\theta))^2}{2\sigma_k^2}\right\} & a < s_k < b \\ 0, & \text{otherwise} \end{cases} \tag{3.2}$$

where $\alpha_{k,i}$ is the probability that the kth pixel is anomalous given the ith hypothesis, σ_k is the standard deviation of the Gaussian distribution for the kth pixel, and a and b are the limits of the range ambiguity interval ($r_{amb} = b - a$) of the LADAR.

A range image of the scene in Figure 3.1 corrupted by coherent or linear mode direct detection LADAR noise is shown in Figure 3.2. Note the presence of the uniformly distributed anomalous pixels scattered through the image. The probability of anomaly is $\alpha_k = 0.15, \forall k$, and the range swath is very short to show the uniform distribution of the anomalies. The image colormap is somewhat truncated so as to preserve the appearance of the scene, resulting in the saturation of some of the anomalies.

3.2.2 Gaussian Noise with Anomalies- Geiger mode LADAR

Due to the nature of Geiger-mode LADAR, it has a somewhat different pdf associated with its anomalous pixels as opposed to either coherent LADAR or linear mode direct detection [19, 21, 21, 23]. We develop a model for it using range images only.

We use the models proposed in [21, 23] for the detection statistics of Geiger mode LADAR. Every range interval has some probability that a photon will enter the APD at the associated time, given that the APD is still looking for photons. Photon sources include laser reflection from the target, background light (e.g., from the sun), and dark current inside the APD.

We assume that the Geiger-mode APDs fire with the arrival of a single photon [21, 23]. As a result, under a Poisson model [23], the probability of at least one photon arriving at a given bin is

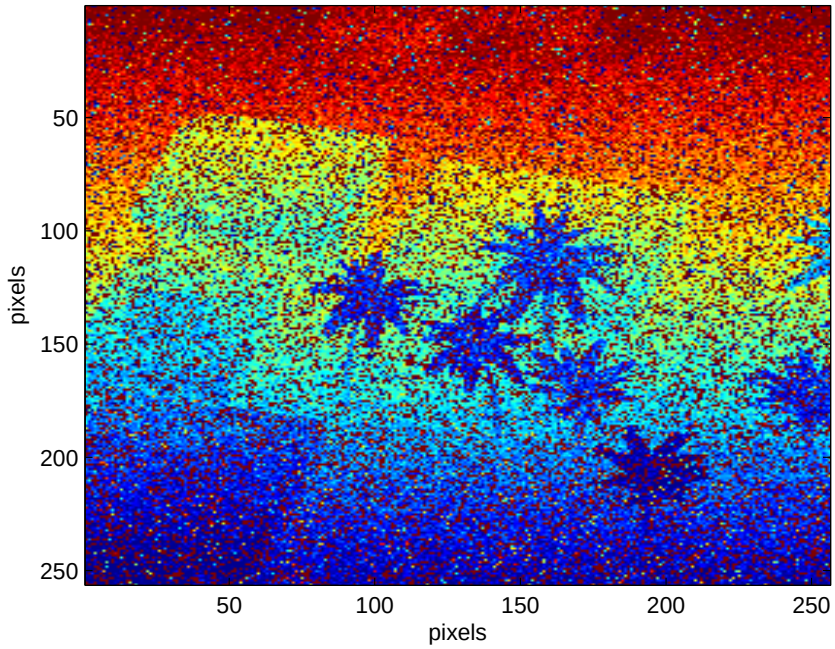

Figure 3.2: Scene corrupted by mixture of Gaussian and uniform distributions, according to the specifications of the simple coherent LADAR noise model. The probability of anomaly is 0.15.

$1-e^{-\lambda}$, where λ depends on a variety of physical parameters and is $\lambda_0\tau_b$ for no target and $\lambda_t+\lambda_0\tau_b$ for a target reflection present in a given range bin, where τ_b is the time length of each bin. Since if a photon is detected at a closer range bin no detection can occur at a farther bin, the probability of a detection in a given range bin is the product of the probability that at least one photon arrives with the probability that no earlier range bin has had a detection.Thus, the probability that the true target is detected, where $\overline{s}$ is the true range of the target, is [21]

$$\rho_k = \exp\left(-\frac{2\overline{s}_k}{c}\lambda_0\right)\left[1-\exp\left(-\lambda_t-\lambda_0\tau_b\right)\right] \tag{3.3}$$

and the probability of a false detection at a bin closer to the sensor than the target is [21]

$$P_{fa,1}(r) = \exp\left(-\frac{2r}{c}\lambda_0\right)\left[1-\exp\left(-\lambda_0\right)\right], a < r < \overline{s}_k \tag{3.4}$$

Using this, the probability of a false detection in a bin at r, behind the target is

$$P_{fa,2}(r) = \left(1 - P_d \frac{\exp\left(-\lambda_t - \lambda_0 \tau_b\right)}{\left[1 - \exp\left(-\lambda_t - \lambda_0 \tau_b\right)\right]}\right) \exp\left(-\frac{2(r - \overline{s}_k)}{c} \lambda_0\right) \times \left[1 - \exp\left(-\lambda_0 \tau_b\right)\right], \overline{s}_k < r < b \tag{3.5}$$

where $[a, b]$ is the range swath of interest to the sensor.

Extrapolating from the detection statistics results in [21, 23], we use these equations to determine the probabilities of anomaly and the anomalous measurement distribution. We extrapolate to infinitely small bins. This is accurate assuming that the bin length is small compared to the gate length τ_g, that is, $\tau_b \ll \tau_g$. The pdf of the anomalies occurring closer and farther than the target are thus

$$p_a(s_k) = \frac{P_{a1}^{(k)} \hat{p}(s_k, \overline{s}_k) + P_{a2}^{(k)} \check{p}(s_k, \overline{s_k})}{P_{a1}^{(k)} + P_{a2}^{(k)}}, \quad a < s_k < b \tag{3.6}$$

where from (3.4) and (3.5) respectively,

$$\hat{p}(s, \overline{s}) = \begin{cases} \gamma_1 \exp(-\beta s) & a < s < \overline{s} \\ 0 & \text{otherwise} \end{cases} \tag{3.7}$$

$$\check{p}(s, \overline{s}) = \begin{cases} \gamma_2 \exp(-\beta s) & \overline{s} < s < b \\ 0 & \text{otherwise} \end{cases}$$

and

$$\beta = 2\lambda_0 / c, \quad \gamma_1 = \frac{\beta}{e^{-\beta a} - e^{-\beta \overline{s}_k}}, \quad \gamma_2 = \frac{\beta}{e^{-\beta \overline{s}_k} - e^{-\beta b}} \tag{3.8}$$

The probabilities of the different components are then

$$\alpha_{k,i}^{(1)} = 1 - \exp\left(-\beta \overline{s}_{k,i}\right) \tag{3.9}$$

$$\alpha_{k,i}^{(2)} \alpha_{k,i}^{(2)} = \exp\left(-\beta \overline{s}_{k,i}\right) \exp\left(-\lambda_t\right) \left[1 - \exp\left(-\beta(b - \overline{s}_{k,i})\right)\right] \tag{3.10}$$

Let the total probability of anomaly then be denoted by

$$\alpha_{k,i} = \alpha_{k,i}^{(1)} + \alpha_{k,i}^{(2)} \tag{3.11}$$

The location of the detection varies based on the length of the LADAR pulse, since photons from the reflection can arrive anytime during the returned pulse [21]. In addition, random local oscillator time jitter causes additional Gaussian range noise [26].

As a result, the pdf associated with Geiger mode LADAR is a mixture of a Gaussian pdf, two truncated exponential pdfs, and dropouts or random deletions. A dropout occurs when no bin observes a photon. We set the pixel value to the maximum range b when a dropout, or deletion, occurs. Note that if multiple possible targets are present, more exponential pdfs are required, and multiple Gaussian distributions may be present. We do not discuss this here since the extension using the mixture pdf model is clear. The complete pixel pdf is given by

$$\begin{aligned} p\left(s_k|H_i\right) &= \alpha_{k,i}^{(1)}\hat{p}(s_k) + \alpha_{k,i}^{(2)}\check{p}(s_k) \\ &+ \frac{\rho_{k,i}}{\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(s_k - \overline{s}_{k,i}(\underline{\theta}))^2}{2\sigma_k^2}\right\} + P_{del}^{(k,i)}\delta(s_k - b), \quad a < s_k \leq b \end{aligned} \tag{3.12}$$

where the probability of deletion is

$$P_{del}^{(k,i)} = 1 - \alpha_{k,i} - \rho_{k,i} \tag{3.13}$$

The Gaussian noise tends to be correlated with all other pixels in a particular image [26], although the number of pixels in a Geiger-mode Flash LADAR tends to be low. Although it is possible to use the correlation model for our performance prediction methods, we use the IID pixel model due to a lack of simple models for the correlation in the literature. This assumption will result in pessimistic performance prediction.

Figure 3.3 shows the scene of Figure 3.1 corrupted by Gaussian noise with Geiger-mode anomalies according to the pdf (3.12). Note the higher concentration of anomalous pixels in the region where the scene has a longer range, due to the fact that early scene objects limit the length of

the high likelihood region of the anomaly pdf. The parameters are chosen such that the probability of anomaly without any object in the scene is 0.15 and the probability of pixel deletion is 0.05. The actual probabilities vary from pixel to pixel based on the associated ranges. The atmospheric attenuation was also exaggerated so as to show its effect of increasing the anomaly likelihood at greater ranges.

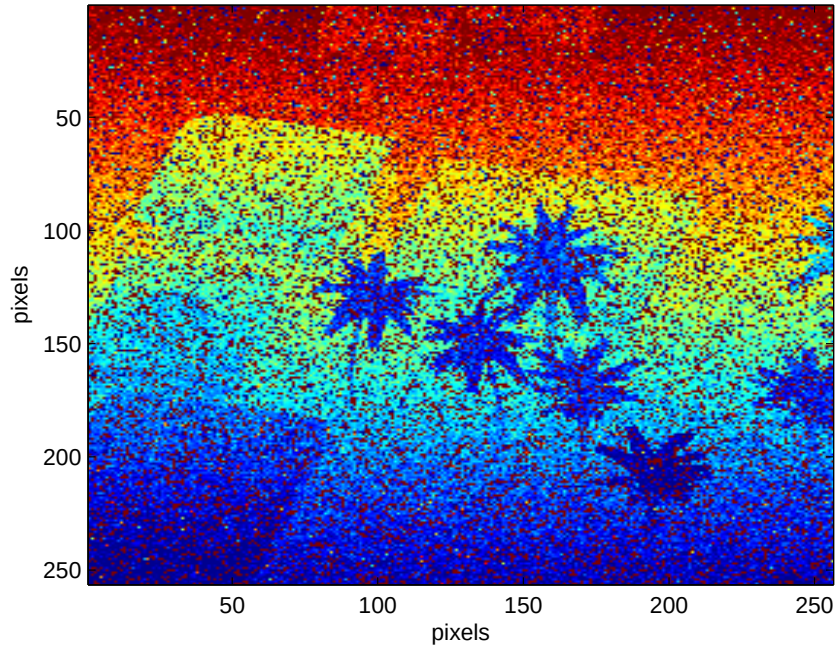

Figure 3.3: Geiger mode LADAR noise including anomalies, deletions, and Gaussian noise. Probability of anomaly 0.15 and probability of deletion 0.05, both defined in the absence of a target.

3.2.3 Gaussian Noise with Random Deletions

Assume as before that the noise for each pixel is independent, with the kth having a Gaussian non-anomalous pdf under H_i and nuisance parameters with mean 0 and covariance $\mathbf{\Sigma}_{k,i}$ and a probability of deletion $\alpha_{k,i}(\underline{\theta})$. Let $\mathcal{A} = \mathcal{R}^K \subset \mathcal{C}^K$, where $\mathcal{R}^K$ is the set of real K-dimensional vectors. The use of $\mathcal{C}^K$ is arbitrary, and is only needed to allow pixels to take on values in sets disjoint from $\mathcal{A}$. Given hypothesis H_i, the kth pixel is deleted with probability $\alpha_{k,i}(\underline{\theta})$. If the pixel is not deleted, it

has the distribution

$$p(\underline{s}_k|H_i) = \frac{1}{\sqrt{(2\pi)^K|\mathbf{\Sigma}_{k,i}|}} \exp\left\{-\frac{1}{2}\left(\underline{s}_k - \overline{\underline{s}}_{k,i}\right)^T \mathbf{\Sigma}_{k,i}^{-1}\left(\underline{s}_k - \overline{\underline{s}}_{k,i}\right)\right\} \tag{3.14}$$

To obtain a pdf for this model, let the deleted pixels be modeled as pixels which take on values from a uniform distribution $u(\underline{s}_k) = 1/|\mathcal{B}|$ on the set $\mathcal{B}$, which is any set in $\mathcal{C}^K$ disjoint from $\mathcal{A}$ thus eliminating the possibility of confusion. The notation $|\mathcal{B}|$ indicates the volume of $\mathcal{B}$. The sets $\mathcal{A}, \mathcal{B}$ must not change with hypothesis. This gives pixel pdfs of the form

$$p\left(\underline{s}_k\right) = \begin{cases} \frac{1-\alpha_{k,i}(\underline{\theta})}{\sqrt{(2\pi)^K|\mathbf{\Sigma}_{k,i}|}} \exp\left\{-\frac{1}{2}\left(\underline{s}_k - \overline{\underline{s}}_{k,i}\right)^T \mathbf{\Sigma}_{k,i}^{-1}\left(\underline{s}_k - \overline{\underline{s}}_{k,i}\right)\right\} \\ \qquad\qquad \underline{s}_k \in \mathcal{A} \\ \alpha_{k,i}(\underline{\theta}) \quad \underline{s}_k \in \mathcal{B} \end{cases} \tag{3.15}$$

By independence, this gives an image pdf of

$$p\left(\mathcal{I}|H_i, \underline{\theta}\right) = \prod_{k=1}^{N} p\left(\underline{s}_k|H_i, \underline{\theta}\right)$$

As the set $\mathcal{B}$ is the same for every hypothesis, it has no effect on the likelihood ratio. As a result, the volume $|\mathcal{B}|$ is arbitrary. For convenience, we take the limit as $|\mathcal{B}| \to \infty$, thus making the likelihood of any realization of a deleted pixel infinitesimal without actually affecting the probability of deletion.

This noise model allows the incorporation of finite beamwidth uncertainty to some degree, in that a 3-D covariance can be selected to approximate the distribution derived using a more accurate technique. This, and the ability to model point clouds obtained from very large numbers of individual images, are the primary reasons we introduce separate point cloud models instead of converting point clouds to range images.

3.2.4 Finite beamwidth effects

For each of the noise models considered so far, the only noise has been range noise. This creates the false impression that the outline of the target as measured by the LADAR is very clean except

for discretization effects. This is frequently not the case, however, especially at large ranges, due to finite beamwidth effects.

As discussed above, each receiver in a LADAR has a finite beamwidth, and thus can receive laser reflections from any point in a small portion of the field of view. The exact point off of which a received photon has reflected is random, with a probability density $g_k(x, y)$ related to the power cross-section of the LADAR beam. This can be converted into a pixel (range) pdf $g_r(\underline{s}_k)$. Regardless of the point observed, the range measurement is corrupted by Gaussian noise with variance σ_k^2. In order to implement this, we discretize the range pdf by dividing the cross-section into regions with roughly the same range values and converting the associated probability into a range probability density function. This is then convolved with the Gaussian noise pdf. Alternatively, the range pdf could be obtained directly, or a Gaussian mixture pdf could be fit to it.

For our model, we use a Gaussian mixture model, for either range images or point clouds. This is to enable the use of the Gaussian and mixture pdf results to be derived later. As a result, the non-anomalous range distribution for K dimensional pixels is

$$p_{na}(\underline{s}_k|H_i) = \sum_{r=1}^{R_m^{(k,i)}} \frac{w_{r,k}}{\sqrt{(2\pi)^K |\mathbf{\Sigma}_{r,k}|^2}} e^{-\frac{1}{2}(\underline{s}_k - \underline{\mu}_{r,k})^T \mathbf{\Sigma}_{r,k}^{-1} (\underline{s}_k - \underline{\mu}_{r,k})} \tag{3.16}$$

where $w_{r,k}$ is the associated probability for each portion, $\underline{\mu}_{r,k}$ is the mean, and $\mathbf{\Sigma}_{r,k}$ is the variance. This distribution can be used with the appropriate anomaly pdf p_a (deletion, uniform, or exponential) and probability of anomaly, giving

$$p(\underline{s}_k) = \alpha_{k,i} p_a(\underline{s}_k|k) + (1 - \alpha_{k,i}) p_{na}(\underline{s}_k|k) \tag{3.17}$$

where $\alpha_{k,i}$ is the probability of anomaly as before. For point clouds, p_a should correspond to the deletion model.

As an example, Figure 3.4 shows noisy images corrupted by linear mode direct detection LADAR finite beamwidth effects and anomalous distributions. The Gaussian noise is IID regardless of the value of r. The half power beamwidth is approximately 2 pixels wide on the focal plane array in both cases.

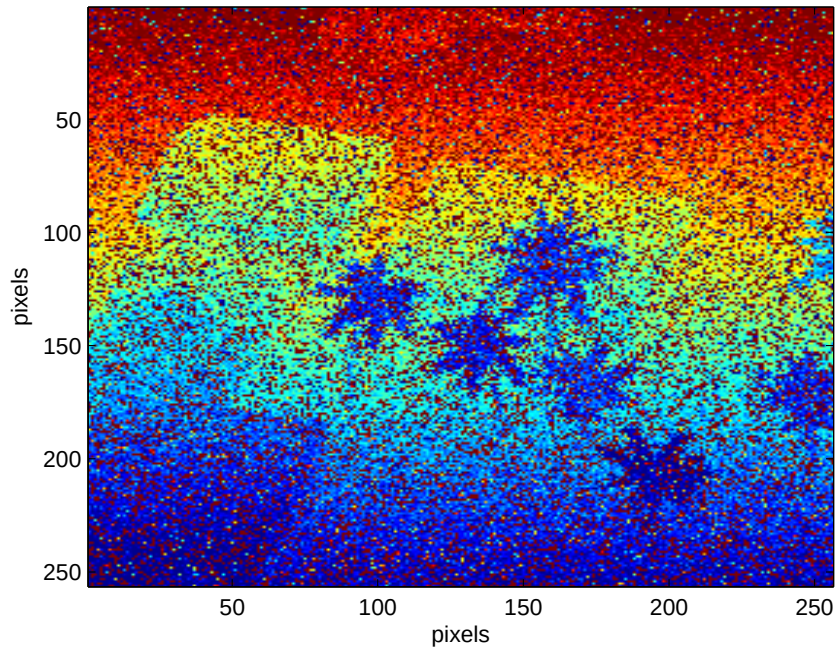

Figure 3.4: Linear mode LADAR noise including anomalies, deletions, finite beamwidth effects, and Gaussian noise. Probability of anomaly 0.15 and probability of deletion 0.05, both defined in the absence of a target.

Chapter 4

Bayesian Classification Performance Prediction

4.1 Introduction

It can be shown that the Bayesian likelihood ratio test is the classifier with the smallest probability of error [13]. We therefore attempt to predict the asymptotic performance of the Bayesian classifier under nuisance parameters, as the image noise variance approaches zero. The classifier is given an image of a target of unknown class and pose in a scene, and its task is to decide which of a set of target classes the target is a member. The image is corrupted by noise of known parameters, and perfect noiseless images of the targets in the scene for every value of the nuisance parameters are assumed available. The performance of this classifier will then be an asymptotic upper bound on the performance of any classification algorithm.

We first review the prediction method presented in [1]. We then derive two more methods of prediction. The first of these methods is an extension of the method presented in [1, 2] with the Laplace approximation replaced by numeric integration. The second method proceeds along a similar line to the first, but modifies the numeric integrand in such a way as to remove a bias inherent in the first method. This method requires the evaluation of an expectation, which must be analytically obtainable and is so for every noise type discussed in this thesis. Due to the nature of

the prediction methods, an added diagnostic benefit is that the methods compute the approximate performance in such a way that the effects of the nuisance parameters on performance are separately computed from the effect of the minimum statistical distance between the targets, and the effects that individual nuisance parameters have on performance are clearly distinguishable.

The remainder of this chapter is organized as follows. Section 4.2 introduces the Bayesian classifier under nuisance parameters. Section 4.3 presents the commonly used method for predicting the asymptotic performance of the classifier. Sections 4.4 and 4.5 describe two new algorithms for predicting asymptotic performance. A method of gradually increasing the prediction accuracy, especially when in the presence of target symmetries, at the cost of increasing computational complexity is introduced in Section 4.7.1 Section 4.7.2 extends the binary classification results to the M-ary classification scenario. An extension for the case where the classifier uses an incorrect noise model is described in Section 4.7.3.

4.2 General Bayesian Hypothesis Testing under Nuisance Parameters

Consider a binary classification scenario. Let the hypotheses that the observed image $\mathcal{I}$ is an image of Target 0 or Target 1 on an identical background be denoted by H_0 and H_1, respectively. Furthermore, let there be an $(m \times 1)$ vector $\underline{\theta}$ of nuisance parameters, such as pose and location, that are unknown to the classifier but have an effect on the distribution of the image. Let the probability density functions (pdf) of the observed image given hypothesis H_i and nuisance parameters $\underline{\theta}$ be denoted by $p(\mathcal{I}|H_i, \underline{\theta})$. Let $f_i(\underline{\theta})$, $i \in \{0, 1\}$ be the priors on the nuisance parameters for the two hypotheses H_0 and H_1, respectively. Then, the probability density function of image $\mathcal{I}$ given hypothesis H_i is obtained by integrating out the nuisance parameters $\underline{\theta}$:

$$p(\mathcal{I}|H_i) = \int_S p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \tag{4.1}$$

Following Bayesian hypothesis testing, the optimal classifier maximizes the posterior probability by comparing the ratio of the probabilities given each hypothesis to a threshold.

$$L(\mathcal{I}) = \log \frac{p(\mathcal{I}|H_1)}{p(\mathcal{I}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \log \frac{P(H_0)}{P(H_1)} \equiv \log \nu \tag{4.2}$$

where $P(H_i)$, $i = 1, 2$ are the prior probabilities of each target occurring in the scene, and $L(\mathcal{I})$ is the loglikelihood ratio.

Predicting the performance of this optimal classifier involves finding the distribution of the log-likelihood ratio $L(\mathcal{I})$. Finding this exactly is typically intractable, necessitating either Monte Carlo methods or approximations. We follow the latter method. The final goal is to find the probabilities of error $P(D_k|H_i),\ i \neq j$ where D_k denotes the event that the classifier has made the decision that the observed image belongs to H_k.

4.3 Method of Grenander et. al.

The method of Grenander et. al. in [1] and others [2, 6, 8] starts by interpreting the pdf of the image as an image with multi-dimensional pixel values corrupted by noise, such as additive Gaussian noise. Let $\underline{I}$ be the vector of pixel values in the image, and $\underline{\overline{I}}_i$ the similarly vectorized underlying noiseless image. K-dimensional pixels are treated as sets of K pixels for this representation. Then, the model is that

$$\underline{I} = \underline{\overline{I}}_i + \underline{n} \tag{4.3}$$

where $\underline{n}$ is a noise vector. Note that no assumption of independence between pixels is made. As a result, the analysis in this and subsequent sections also applies to image models containing multi-dimensional pixels, as is the case for 3-D LADAR point clouds. This is because (for example) a K dimensional model with N pixels is equivalent to a model of KN pixels, with each coordinate of the actual K dimensional pixel being assigned to an equivalent scalar pixel. Let $\mathbf{P} = E[\underline{n}\underline{n}']$ be the

$(KN \times KN)$ matrix giving the expectation of the products of the elements of the noise vector. Next, let the KN eigenvalues of this matrix be given by $\lambda_p, p = 1, ..., KN$. Furthermore, let the noise be parameterized by a variable σ, such that $\lambda_p = \lambda_p(\sigma)$, in such a way that $\lim_{\sigma \to 0} \lambda_p(\sigma) = 0, \forall n$. This parameter will be used exclusively for proving results for the asymptotic behavior of the prediction methods to be presented, and does not imply any loss of generality. As a result, the power of the noise goes to zero as $\sigma \to 0$. Assume further that the ratios of the expectations to the associated standard deviations all approach zero as $\sigma \to \infty$, that is, $\lim_{\sigma \to \infty} \frac{E[\underline{n}(\sigma)]}{\sqrt{\det \mathbf{P}}} = 0$. This guarantees that the probability of error decays to 0.5 as σ increases.

Let the noiseless model images for hypothesis H_i and pose $\underline{\theta}$ be denoted as $\overline{\mathcal{I}}_i(\underline{\theta})$. Let the noiseless truth image for hypothesis H_0, with nuisance parameters $\underline{\theta} = \underline{\theta}_{true}$ be denoted by $\overline{\mathcal{I}}_0^* = \overline{\mathcal{I}}_0(\underline{\theta}_{true})$. The noiseless model images are the images that would be the truth image if the true hypothesis were H_i and the true pose were $\underline{\theta}$. By "noiseless," we mean the image such that

$$\overline{\mathcal{I}}_i(\underline{\theta}) = \arg \max_{\mathcal{I}} p(\mathcal{I}|H_i, \underline{\theta}) \tag{4.4}$$

It is desired to predict the performance in such a way that the predicted performance is asymptotically equal to the true performance as the pixel noise variances all approach zero. This is because as the noise level decreases and becomes small, the probabilities of error become quite small, necessitating a highly accurate prediction in order to be able to preserve meaningful relative accuracy. In contrast to this, errors are more tolerable for high noise, since the probabilities of error are on the order of 0.5, and in most cases it is sufficient to know that the performance is poor, further decreasing the need for high accuracy for this noise variance region. For this type of prediction, we employ the concept of asymptotic equality [1]. We use the notation $x(\sigma) \sim y(\sigma)$ if x and y are asymptotically equal as $\sigma \to 0$, that is, $\lim_{\sigma \to 0} \frac{x(\sigma)}{y(\sigma)} = 1$. This indicates that the error, or $x - y$, also converges to 0 as desired. It can be shown that $x(\sigma) \sim y(\sigma)$ implies $\log(x(\sigma)) \sim \log(y(\sigma))$ [1].

The authors of [1] prove that the prediction method they propose is accurate as $\sigma \to 0$ (i.e., the predicted probability of correct classification is asymptotically equal to the true value) provided that consistency of inference is enforced. Consistency of inference is achieved when the noiseless image $\overline{\mathcal{I}}_i(\underline{\theta})$ is more likely given $H_i, \underline{\theta}$ than any other image for all σ. Since the authors of [1] consider

only zero-mean Gaussian noise, they use the corresponding special case and define consistency of inference to be

$$\overline{\mathcal{I}}_i(\underline{\theta}_1) \neq \overline{\mathcal{I}}_k(\underline{\theta}_2), \quad \forall\, i, j, \underline{\theta}_1, \underline{\theta}_2 \; s.t. i \neq j \; or \; \underline{\theta}_1 \neq \underline{\theta}_2 \tag{4.5}$$

For our analysis, a more general result is required. We thus define consistency of inference to hold when

$$[i, \underline{\theta}_i] = \arg\max_{j, \underline{\theta}_k} p(\overline{\mathcal{I}}_i(\underline{\theta}_i)|H_k, \underline{\theta}_k, \sigma), \forall i, \underline{\theta}_i, \sigma \tag{4.6}$$

which clearly implies Grenander's definition in the case of Gaussian noise, for which is was originally defined, although it is slightly more restrictive. When this applies, as $\sigma \to 0$, the support of $p(\mathcal{I}|H_i, \underline{\theta})$ as a function of i and $\underline{\theta}$ converges to the true values for the distribution of $\mathcal{I}$. Consistency of inference should never be an issue with real targets, due to the nature of pixel noise.

The authors of [1] then proceed to derive an asymptotically accurate approximation to the log-likelihood ratio test, with the goal of obtaining a performance approximation that can be computed in an efficient manner. First, asymptotically approximate the integrands of (4.1) for purposes of performance prediction to be

$$p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) \sim p(\mathcal{I}|H_i, \underline{\theta}_i^*) \frac{p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) f_i(\underline{\theta})}{p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}_i^*)} \tag{4.7}$$

where

$$\underline{\theta}_i^* = \arg\max_{\underline{\theta}} p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) f_i(\underline{\theta}) \tag{4.8}$$

since for the asymptotic case the likelihood ratio is dominated by the performance when the targets are closest as measured by the likelihood function. Using this, consistency of inference, and (4.4), it is clear that $\underline{\theta}_0^* = \underline{\theta}_{true}$.

This gives the asymptotic approximation to the Bayesian integrals (4.1) to be

$$\int_S p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \sim \frac{p(\mathcal{I}|H_i, \underline{\theta}_i^*)}{p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}_i^*)} \int_S p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \tag{4.9}$$

4.3.1 Approximation of Bayesian Integral using Laplace's method

Grenander proposes to approximate the integral in the right hand side of (4.9) using Laplace's method. This method applies for the asymptotic case as $\sigma \to 0$. In essence, Laplace's method approximates the function $\log p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) f_i(\underline{\theta})$ as a parabola centered on its maximum point $\underline{\theta}_i^*$, with second derivatives determined by the second derivatives of the function at $\underline{\theta}_i^*$. This technique is commonly used for obtaining asymptotic approximations of integrals or moments [1, 33, 34].

Using Laplace's method, it can be shown [1] that under certain regularity conditions on the likelihood function, that

$$\int_S p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \sim (2\pi)^{m/2} \sqrt{\frac{(2\sigma^2)^m}{\det\left(\ddot{E}_i(\overline{\mathcal{I}}_0^*, \underline{\theta}_i^*, \sigma)\right)}} \tag{4.10}$$

where

$$E_i(\mathcal{I}, \underline{\theta}, \sigma) = -2\sigma^2 \left(\log p(\mathcal{I}|H_i, \underline{\theta}) + \log f_i(\underline{\theta})\right) + c_e \tag{4.11}$$

where c_e is a constant, E_i is the Gibb's energy function associated with the likelihood, and $\ddot{E}_i$ denotes the Hessian matrix of E_i as a function of $\underline{\theta}$. Interestingly, the Cramer-Rao bound on $\underline{\theta}$ is $\frac{\ddot{E}_i}{2\sigma^2}$, which appears in the expression for the approximate integral. This indicates that, as expected, the accuracy of estimating the nuisance parameters has a direct effect on classifier performance.

Substituting the expression in (4.10) into (4.9) and then into the likelihood ratio (4.2) and taking the logarithm, the following pseudo loglikelihood ratio test results.

$$L'(\mathcal{I}) = \log \frac{p(\mathcal{I}|H_1, \underline{\theta}_1^*)}{p(\mathcal{I}|H_0, \underline{\theta}_0^*)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log \nu' \tag{4.12}$$

where

$$\nu' = \nu \sqrt{\frac{\det\left(\ddot{E}_1(\overline{\mathcal{I}}_0^*, \underline{\theta}_1^*, \sigma)\right)}{\det\left(\ddot{E}_0(\overline{\mathcal{I}}_0^*, \underline{\theta}_0^*, \sigma)\right)}} \tag{4.13}$$

The authors of [1] then proceed to predict the distribution of this test for IID Gaussian pixel

noise. They show that $L'(\mathcal{I})$ is Gaussian distributed. Using this, it is a straightforward Q-function evaluation to obtain the probability of error. We will describe the computation of the predicted probability of error in Section 4.6.

4.4 Method of Noiseless Integration: Eliminate Laplace Approximation

While the method of Grenander is asymptotically correct, it is more important to be able to predict the performance for larger noise powers. In particular, for the Laplace approximation of the integrand as a Gaussian pulse to be valid, the noise level must be very small, and the resulting probability of correct classification very high. Consider the case of two ground targets shown in Figures 4.1 and 4.2, with a nuisance parameter θ of horizontal angular pose. As an example showing the non-Gaussian shape of the integrand for reasonable noise levels, a plot of $\frac{p(\overline{\mathcal{I}}_0^*|H_1,\theta)f_1(\theta)}{p(\overline{\mathcal{I}}_0^*|H_1,\theta_1^*)}$ (the approximating integrand in (4.9)) as a function of target horizontal pose angle θ for Gaussian noise and a low noise level (SNR ≈ 44dB, where the signal power is measured as the root sum squared error between $\overline{\mathcal{I}}_0^*$ and $\overline{\mathcal{I}}_1(\theta_1^*)$) resulting in $P_{cc} = .9995$ is shown in Figure 4.3. In addition, a plot for the same integrand for very high noise (SNR ≈ -20dB), resulting in performance near 50% is shown in Figure 4.4, also showing the poorness of the Gaussian approximation for high noise. Note in the plot that the curve does not decay to zero, and the two tails do not decay to the same value. The result given by the Laplace approximation will vary greatly depending on the choice of the points to sample in order to obtain a value for the second derivative of the logarithm at the maximum point, since the true second derivative is obviously inadequate. Since the Gaussian approximation is not as accurate as may be desired for reasonable noise levels, we seek to remove this approximation. Noiseless LADAR images can be generated for arbitrary nuisance parameters quickly using available software, allowing for the direct evaluation of the required conditional pdfs. Hence, the approximating integral on the right side of (4.9) can be evaluated numerically. We refer to this method as the Noiseless Integration Method (NIM).

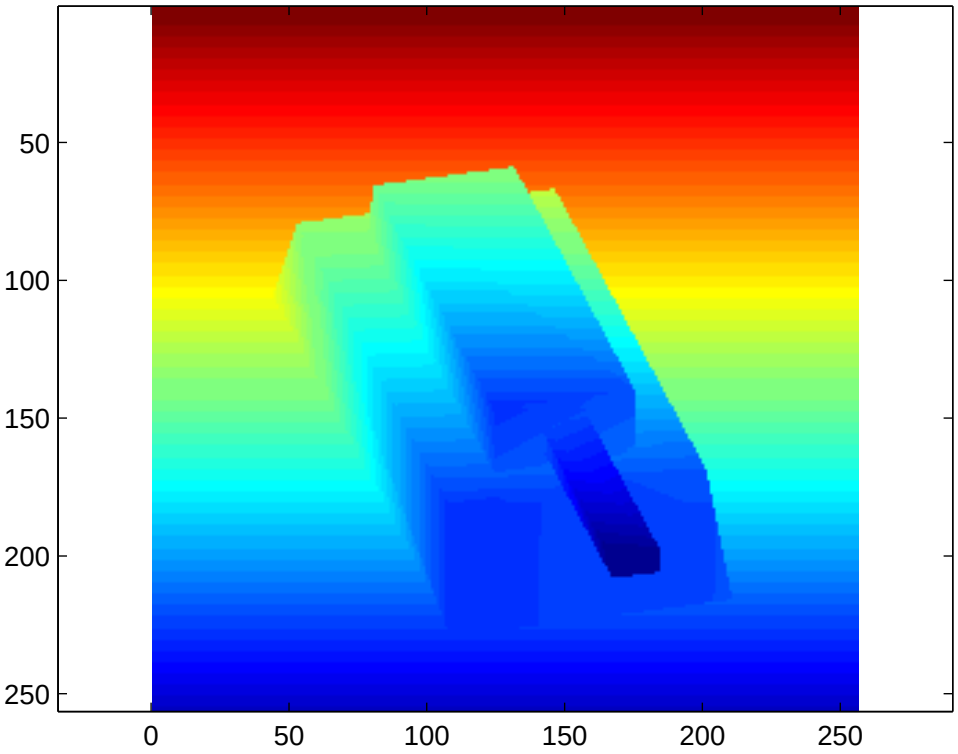

Figure 4.1: View of Tank CAD model from simulated LADAR sensor location.

When the integrals are evaluated, the approximate likelihood test becomes

$$L(\mathcal{I}) \sim \frac{p(\mathcal{I}|H_1, \underline{\theta}_1^*) p(\overline{\mathcal{I}}_0^*|H_0, \underline{\theta}_0^*) \int_S p(\overline{\mathcal{I}}_0^*|H_1, \underline{\theta}) f_1(\underline{\theta}) d\underline{\theta}}{p(\mathcal{I}|H_0, \underline{\theta}_0^*) p(\overline{\mathcal{I}}_0^*|H_1, \underline{\theta}_1^*) \int_S p(\overline{\mathcal{I}}_0^*|H_0, \underline{\theta}) f_0(\underline{\theta}) d\underline{\theta}} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \nu \tag{4.14}$$

which is again equivalent to a threshold adjustment

$$L' = \log \frac{p(\mathcal{I}|H_1, \underline{\theta}_1^*)}{p(\mathcal{I}|H_0, \underline{\theta}_0^*)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log \nu'$$

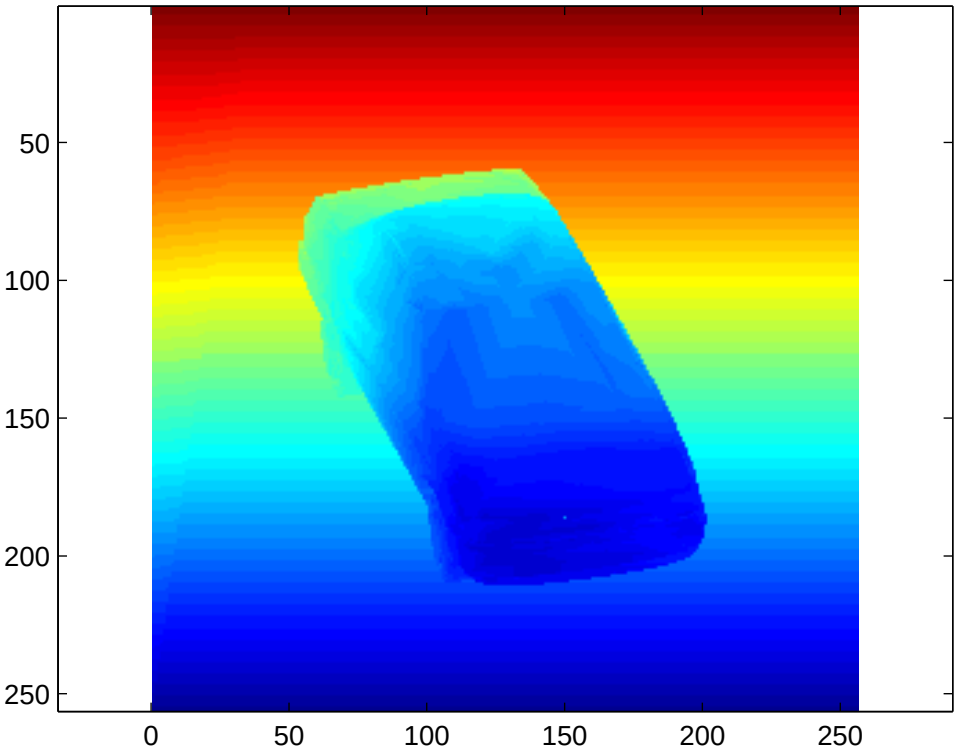

Figure 4.2: View of Sedan CAD model from simulated LADAR sensor location.

where now

$$\nu' = \nu \frac{p(\overline{\mathcal{I}}_0^*|H_1, \underline{\theta}_1^*) \int_S p(\overline{\mathcal{I}}_0^*|H_0, \underline{\theta}) f_0(\underline{\theta}) d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*|H_0, \underline{\theta}_0^*) \int_S p(\overline{\mathcal{I}}_0^*|H_1, \underline{\theta}) f_1(\underline{\theta}) d\underline{\theta}} \tag{4.15}$$

4.5 Unbiased Integration Method

Although the method described in Section 4.4 is asymptotically correct, the approximation of the Bayesian integrand by the integrand when $\mathcal{I} = \overline{\mathcal{I}}_0^*$ (Equation (4.14) is biased in general, since $E[f(x)] \neq f(E[x])$ in general. We desire to continue, however, with the threshold adjustment approximation technique as that used by Grenander et. al. and by the Method of Noiseless Integration because the reduction of the approximation to a mere change in threshold for the pseudo-log-likelihood ratio (4.12) is particularly attractive due to its ease of implementation, as will be discussed in Chapter 5. We therefore seek to derive an improved approximation method.

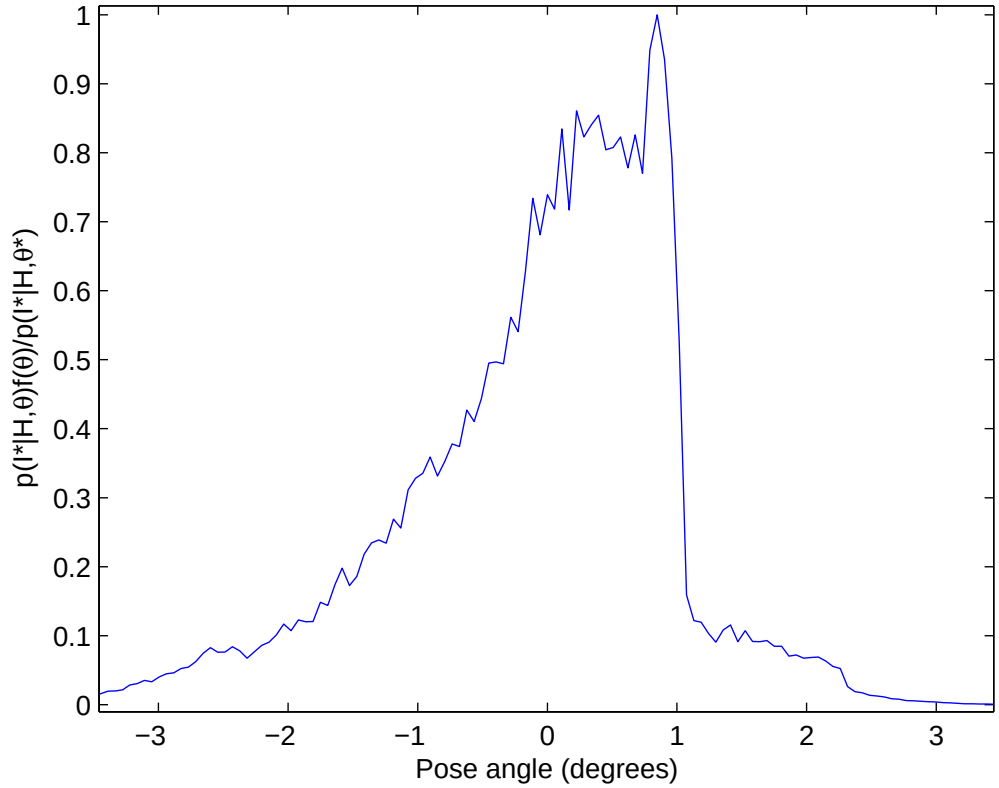

Figure 4.3: Plot of the approximating integrand with respect to angular pose (in degrees) for noise standard deviation of 0.02. Note that the curve has a significantly non-Gaussian shape, indicating that the Laplace approximation is not sufficiently accurate for this relatively low noise level.

As noted above and in [1], the goal of the approximations in Sections 4.3 and 4.4 is to obtain an asymptotically correct approximation to the log-likelihood ratio by adding a constant to the pseudo log-likelihood ratio in (4.12). In other words, a constant ν' is obtained such that

$$\lim_{\sigma \to 0} \log \left(\frac{\nu}{\nu'} \frac{p(\mathcal{I}|H_i, \underline{\theta}_i^*)}{p(\mathcal{I}|H_0, \underline{\theta}_0^*)} \right) - \log \left(\frac{\int_S p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta}}{\int_S p(\mathcal{I}|H_0, \underline{\theta}) f_0(\underline{\theta}) d\underline{\theta}} \right) = 0 \tag{4.16}$$

regardless of $\mathcal{I}$, using the appropriate equivalent expression for asymptotic equality in the log domain. The threshold ν' is not selected, however, based on any considerations regarding the rate at which the error asymptotically converges to zero. Since the approximation in Section 4.4 is biased, it is quite likely that the error is also biased as it converges to zero. This is not desirable, since the goal of the prediction method is to predict the performance for as large a range of σ as possible. As

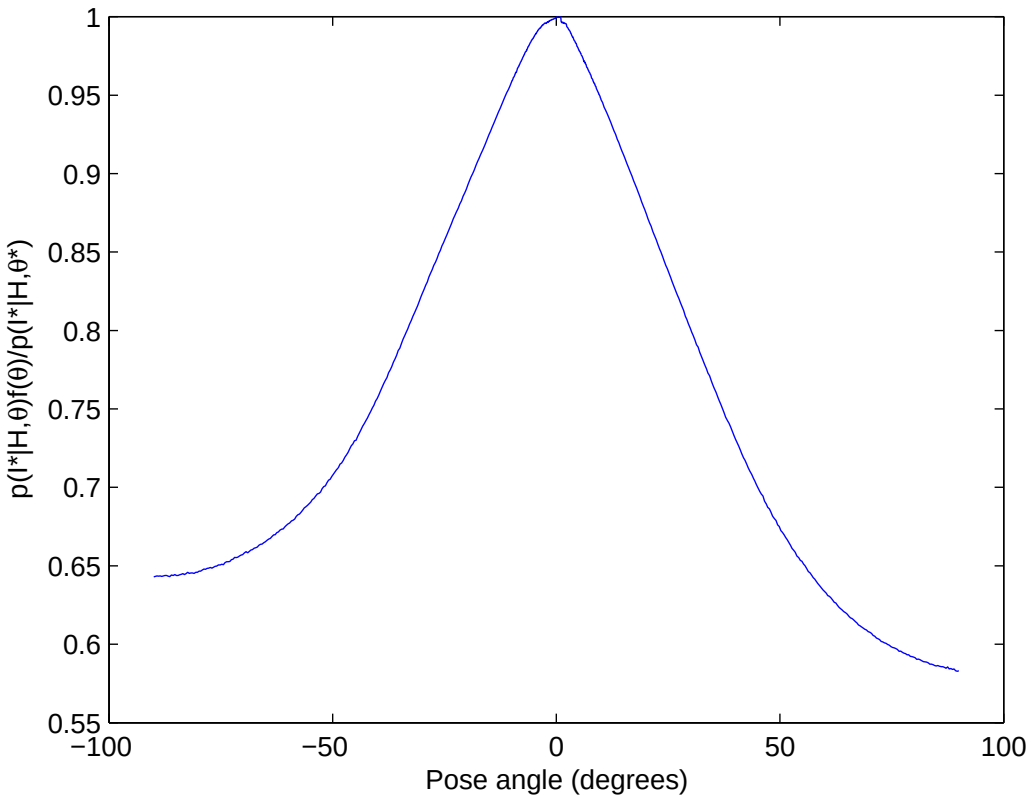

Figure 4.4: Plot of the approximating integrand with respect to angular pose (in degrees) for noise standard deviation of 0.5. Note that the curve has a significantly non-Gaussian shape, indicating that the Laplace approximation is not sufficiently accurate for high noise levels.

a result, we seek approximations that result in errors near zero for as large a value of σ as possible. Since the threshold adjustment ν' has no way of independently modifying the the expectation and variance of the pseudo log-likelihood ratio, and since in the log domain the threshold adjustment is an additive constant, it is reasonable to try to obtain a ν' for which the expected error is zero, considering $\mathcal{I}$ as a random variable. This is reasonable since in the log domain the log-likelihood ratio is of course created by taking the difference between two log-likelihoods.

This is further supported intuitively by the fact that for large numbers of independent pixels, the Central Limit Theorem (with some assumptions on the nature of the pixel pdfs) gives the result that the pseudo log-likelihood ratio is close to Gaussian distributed. Since the true log-likelihood ratio is asymptotically equivalent, it must also be asymptotically nearly Gaussian distributed. Hence, the

distributions should not be heavily skewed, so that the expectation should be near the high likelihood region, thus ensuring that the error incurred should be near zero most of the time if the expected error is zero.

To obtain an expectation of zero, we must have that

$$E\left[\log\left(\frac{\nu}{\nu'}\frac{p(\mathcal{I}|H_i,\underline{\theta}_i^*)}{p(\mathcal{I}|H_0,\underline{\theta}_0^*)}\right)-\log\frac{\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta}}{\int_S p(\mathcal{I}|H_0,\underline{\theta})f_0(\underline{\theta})d\underline{\theta}}\right]=0 \tag{4.17}$$

or

$$\begin{aligned}&E\left[\log(C_i p(\mathcal{I}|H_i,\underline{\theta}_i^*))-\log(\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta})\right]\\&-E\left[\log(C_0 p(\mathcal{I}|H_0,\underline{\theta}_0^*))-\log(\int_S p(\mathcal{I}|H_0,\underline{\theta})f_0(\underline{\theta})d\underline{\theta})\right]=0\end{aligned} \tag{4.18}$$

where $C_i/C_0=\nu/\nu'$. Dividing ν' up in this manner allows us to separate the problem into two by seeking C_0 and C_i such that

$$\begin{aligned}E\left[\log(C_i p(\mathcal{I}|H_i,\underline{\theta}_i^*))-\log(\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta})\right]=0\\E\left[\log(C_0 p(\mathcal{I}|H_0,\underline{\theta}_0^*))-\log(\int_S p(\mathcal{I}|H_0,\underline{\theta})f_0(\underline{\theta})d\underline{\theta})\right]=0\end{aligned} \tag{4.19}$$

which will achieve the overall zero expected error in a completely equivalent manner.

It is very difficult, however, to obtain the expectation of the log of a likelihood integral for most pdfs of interest, due to the nonadditivity of the terms inside the logarithm. This problem is somewhat related to that of computing the distribution of the sum of lognormal random variables, for which exact solutions are intractable [35, 36]. We thus seek to move the expectation inside the integral, so as to be able to move the expectation inside the integral and exploit any pixel independence, thus making the problem tractable.

We thus use a moment matching technique [36] to approximate the Bayesian integral itself instead of its logarithm. That is,

$$E\left[\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta}-Cp(\mathcal{I}|H_i,\underline{\theta}_i^*)\right]=0 \tag{4.20}$$

If the expected error in the likelihood domain is zero and the truth and the approximation are asymptotically equal for every value of $\mathcal{I}$ as desired, the expected error in the log-likelihood domain will be asymptotically small. This approach is equivalent to approximating via moment matching, and is in some ways based on a similar principle as the approach of [36] to the problem of approximating the discrete sum of IID lognormal random variables by a single lognormal.

Solving (4.20) for C gives

$$C = \frac{\int_S E\left[p(\mathcal{I}|H_i, \underline{\theta})\right] f_i(\underline{\theta}) d\underline{\theta}}{E[p(\mathcal{I}|H_i, \underline{\theta}_i^*)]} \tag{4.21}$$

This gives an approximation to the Bayesian integral

$$\int_S p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \sim p(\mathcal{I}|H_i, \underline{\theta}_i^*) \frac{\int_S E[p(\mathcal{I}|H_i, \underline{\theta})] f_i(\underline{\theta}) d\underline{\theta}}{E[p(\mathcal{I}|H_i, \underline{\theta}_i^*)]} \tag{4.22}$$

which the following theorem proves to be asymptotically accurate. We refer to this approximation method as the Unbiased Integration Method (UIM).

Theorem 1 *The approximation in* (4.22) *is asymptotically correct, that is*

$$\int_S p(\mathcal{I}|H_i, \underline{\theta}) f_i(\underline{\theta}) d\underline{\theta} \sim p(\mathcal{I}|H_i, \underline{\theta}_i^*) \frac{\int_S E[p(\mathcal{I}|H_i, \underline{\theta})] f_i(\underline{\theta}) d\underline{\theta}}{E[p(\mathcal{I}|H_i, \underline{\theta}_i^*)]} \tag{4.23}$$

The proof is given in Appendix A.

This approximation is then substituted into the Bayesian likelihood ratio (4.2) as was done in (4.14) resulting in

$$L' = \log \frac{p(\mathcal{I}|H_1, \underline{\theta}_1^*)}{p(\mathcal{I}|H_0, \underline{\theta}_0^*)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log \nu' \tag{4.24}$$

where the threshold adjustment ν' is now given by

$$\nu' = \nu \frac{E\left[p(\mathcal{I}|H_1, \underline{\theta}_1^*)\right] \int_S E\left[p(\mathcal{I}|H_0, \underline{\theta})\right] f_0(\underline{\theta}) d\underline{\theta}}{E\left[p(\mathcal{I}|H_0, \underline{\theta}_0^*)\right] \int_S E\left[p(\mathcal{I}|H_1, \underline{\theta})\right] f_1(\underline{\theta}) d\underline{\theta}} \tag{4.25}$$

The calculation of the expectation in (4.22) will be discussed in Chapter 5 and can be less straightforward than the calculation of the integrands in (4.15), but the resulting prediction is significantly more accurate, especially when the noise level is increased.

As with the method of Grenander, if computational limitations require, the Laplace approximation can be used for this method as well. It can be shown that for Gaussian noise, this gives identical results to Grenander's method.

4.6 Prediction of Probability of Error

Having computed the threshold adjustment ν' for the pseudo loglikelihood ratio test using one of the three prediction methods (4.13),(4.15),(4.38), it remains to compute the probability of error. The pseudo loglikelihood ratio is given by

$$L' = \log \frac{p(\mathcal{I}|H_1, \underline{\theta}_1^*)}{p(\mathcal{I}|H_0, \underline{\theta}_0^*)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log \nu' \tag{4.26}$$

Hence, the predicted probability of error $P(D_1|H_0)$ for this test can be calculated as

$$P(D_1|H_0) \sim \int_{\log \nu'}^{\infty} p(l|H_0) dl$$

so long as the pdf of L', denoted by $p(l|H_0) = p(L' = l|H_0)$, can be derived or numerically calculated. The accuracy of this method will be limited by the accuracy of the numerical integrations performed. This prediction is asymptotically equal to the true probability of error.

4.7 Extensions

In this section, we discuss several extensions of the the Noiseless Integration Method (NIM) and Unbiased Integration Method (UIM) prediction methods. First, a multimodal extension is consid-

ered where multiple "pseudo loglikelihoods" are used to approximate the full Bayesian likelihood ratio test, as opposed to the single pseudo LLR test. This method allows for increased prediction accuracy. Second, an extension is considered that allows the prediction of different types of error probabilities in the M-ary target classification scenario. Finally, we consider the prediction of the performance of Bayesian classifiers derived using incorrect noise models and/or parameters.

4.7.1 Multi-modal Extension of NIM and UIM

Due to the symmetries of some targets, it is common that the likelihood as a function of the nuisance parameters is multi-modal. For example, most civilian vehicles, especially sedans, have a roughly similar appearance from above after a 180 degree rotation. In this case, the likelihood as a function of angular pose is multi-modal. If the noise level is high enough, the contribution to the Bayesian integrals from the secondary peak of the function (located at roughly 180 degrees from the true angle) is significant. The methods developed up to this point would integrate the appropriate integrand over the entire nuisance parameter space and use the result to adjust the likelihood ratio test threshold, using the likelihood ratio at the primary peak of the function. This essentially approximates the integral as highly correlated with the random value of the integrand at the primary peak. This is a quite good approximation when the region for which the integrand is of significant magnitude is solely in the neighborhood of the primary peak. The presence of secondary peaks, however, increases the chance that portions of the integrand with significant magnitude will be slightly uncorrelated. Thus, in order to improve the prediction accuracy, we propose to use the random likelihoods at several nuisance parameter values, instead of only at the primary peak. In other words, for the Unbiased Integration Method, the approximating integral derived above

$$\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta} \sim p(\mathcal{I}|H_i,\underline{\theta}_i^*)\frac{\int_S E[p(\mathcal{I}|H_i,\underline{\theta})]f_i(\underline{\theta})d\underline{\theta}}{E[p(\mathcal{I}|H_i,\underline{\theta}_i^*)]}$$

is replaced with the approximation

$$\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta} \sim \sum_{w=1}^{W_i} p(\mathcal{I}|H_i,\underline{\theta}_i^{(w)})\frac{\int_{S_{w,i}} E[p(\mathcal{I}|H_i,\underline{\theta})]f_i(\underline{\theta})d\underline{\theta}}{E[p(\mathcal{I}|H_i,\underline{\theta}_i^{(w)})]}$$

where $S_{w,i}$ are disjoint sets such that $\underline{\theta}_i^{(w)} \in S_{w,i}$ and $\cup_{w=1}^{W_i} S_{w,i} = S$, and using W_i random samples located at $\underline{\theta}_i^{(w)}$ in the approximation. For asymptotic accuracy, the choice of $\underline{\theta}_i^{(w)}$ should include $\underline{\theta}_i^*$. It is natural that the other points be chosen at the secondary peaks, although this method is not limited by this. The ranges for each partial integral would typically be chosen so that points are associated with the peaks closest to them. The most natural boundaries between sets would lie along the local valleys between the peaks. For asymptoticity, it is required that all integrals not containing $\underline{\theta}_i^*$ approach zero as $\sigma \to 0$. For appropriately chosen $S_{w,i}$ and $\underline{\theta}_i^{(w)}$, the approximation should become increasingly accurate as the W_i increase. It is also clear from the nature of the approximation that assuming $p(\mathcal{I}|H_i, \underline{\theta})$ is a continuous function of $\underline{\theta}$, arbitrary approximation accuracy can be achieved for W_i sufficiently large, at the expense of rapidly increasing computational complexity.

This approximation thus gives an gives an approximate likelihood ratio test of

$$L' = \log\left(\frac{\sum_{w=1}^{W_1} p(\mathcal{I}|H_1, \underline{\theta}_1^{(w)}) \frac{\int_{S_{w,1}} E[p(\mathcal{I}|H_1,\underline{\theta})] f_1(\underline{\theta}) d\underline{\theta}}{E[p(\mathcal{I}|H_1,\underline{\theta}_1^{(w)})]}}{\sum_{w=1}^{W_0} p(\mathcal{I}|H_0, \underline{\theta}_0^{(w)}) \frac{\int_{S_{w,0}} E[p(\mathcal{I}|H_0,\underline{\theta})] f_0(\underline{\theta}) d\underline{\theta}}{E[p(\mathcal{I}|H_0,\underline{\theta}_0^{(w)})]}}\right) \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log \nu \tag{4.27}$$

and the probability of error is found using the pdf of L' as before.

Of course, this multi-modal approach is also applicable to the NIM. As the derivation closely follows that for the UIM, only the end result is given here. The approximating likelihood ratio is given by

$$L' = \log\left(\frac{\sum_{w=1}^{W_1} p(I|H_1, \underline{\theta}_1^{(w)}) \frac{\int_{S_{w,1}} p(\overline{\mathcal{I}}_0^*|H_1,\underline{\theta}) f_1(\underline{\theta}) d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*|H_1,\underline{\theta}_1^{(w)})}}{\sum_{w=1}^{W_0} p(\mathcal{I}|H_0, \underline{\theta}_0^{(w)}) \frac{\int_{S_{w,0}} p(\overline{\mathcal{I}}_0^*|H_0,\underline{\theta}) f_0(\underline{\theta}) d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*|H_0,\underline{\theta}_0^{(w)})}}\right) \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \nu \tag{4.28}$$

Since this multi-modal method requires the computation of the pdf of a function of $W_0 + W_1$ correlated random variables, it is more expensive to compute than the prediction methods using a single approximating random sample. As a result, the W_i should be kept small to preserve compu-

tational efficiency. Typically, the number of random samples used will be limited to the number of distinct peaks in the Bayesian integrand, as determined by the approximate symmetries of the target relative to the nuisance parameters used.

4.7.2 Extension to M-ary Classification

So far, we have only considered the case of binary classification. More realistic ATR scenarios almost always involve classification amongst multiple possible classes. We thus extend the results up to this point to the case of M possible target classes [2].

Let the M classes be denoted by hypotheses H_i, $i = 0, ..., M-1$. The optimum Bayesian classifier is given by

$$i = \arg\max_{i_D} P(\mathcal{I}|H_{i_D})P(H_{i_D}) \tag{4.29}$$

where i is the index of the selected hypothesis.

It is desired to calculate the confusion matrix C of the system. The confusion matrix consists of the probabilities that hypothesis i is selected when in fact hypothesis j is the truth.

$$C_{ij} = P(D_i|H_j), \quad \forall i, j \tag{4.30}$$

where as before D_i indicates that the classifier selected H_i.

To compute C_{ij}, first note that the classifier in (4.29) dictates that D_i will occur if and only if

$$\frac{P(\mathcal{I}|H_i)P(H_i)}{P(\mathcal{I}|H_j)P(H_j)} > 1, \quad \forall j \neq i \tag{4.31}$$

Following the techniques of Sections 4.3, 4.4 or 4.5, the likelihood ratio tests in (4.31) can be approximated after taking the logarithm as

$$L_{ij} = \log p(\mathcal{I}|H_i, \underline{\theta}_i^*) - \log p(\mathcal{I}|H_j, \underline{\theta}_j^*) > \log \nu'_{ij}, \forall j \neq i \tag{4.32}$$

where ν'_{ij} is the appropriate threshold adjustment as dictated by one of the three methods just mentioned.

For notational simplicity, let

$$\underline{L}_i = \begin{bmatrix} L_{i1} & \dots & L_{i,i-1} & L_{i,i+1} & \dots & L_{i,M-1} \end{bmatrix}^T \tag{4.33}$$

be a vector containing each of the log-likelihood ratios $L_{ij}, \forall j \neq i$.

Let $p(\underline{L}_i|H_\ell)$ be the joint pdf of the log-likelihood ratios L_{ij}, $\forall j \neq i$ given that H_ℓ is true. Then, let region $A = \cap_{j=0,j\neq i}^{M-1}[L_{ij} > \nu'_{ij}]$. Then,

$$P(D_i|H_\ell) = \int_A p(\underline{L}_i|H_\ell) d\underline{L}_i \tag{4.34}$$

The most difficult part of this method is finding the joint pdf of the log-likelihood ratios. For Gaussian noise it is straightforward to derive the joint pdf by obtaining the mean and covariance since each log-likelihood ratio is a linear combination of Gaussian variables ([2] and see below). For other types of noise, it is usually possible to approximate the pdf of the log-likelihood ratios as Gaussian for a large enough number of pixels, due to the Central Limit Theorem and independence of each pixel's noise. This approximation simplifies the task, since only the mean and covariance would need to be calculated, probably using numerical techniques.

4.7.3 Extension to Incorrect Classifier Noise Models

Due to the complexity of accurate noise models and the need to estimate noise model parameters, it is frequently the case that an ATR classifier is based on simplified noise models or slightly incorrect noise parameters [15]. To evaluate whether or not the use of the incorrect noise model is tolerable, it is important to be able to predict the performance of the classifier derived optimally for the wrong noise model when it is applied to data using the true noise model. We thus present methods of doing this using the NIM and UIM. Let $p_t(\cdot)$ indicate the true noise model pdf, and $p_c(\cdot)$ the noise model pdf used by the classifier. The priors used are those of the classifier's noise model. Let the threshold computed for this prediction scenario be ν'.

We first consider the Noiseless Integration Method. As the threshold ν' is computed using the noiseless likelihood values, the likelihood functions should be those dictated by the incorrect model,

and the noiseless image be determined by the true noise model. That is, following 4.15,

$$\nu' = \nu \frac{p_c(\overline{\mathcal{I}}^*_{0,t}|H_1, \underline{\theta}^*_1) \int_S p_c(\overline{\mathcal{I}}^*_{0,t}|H_0, \underline{\theta}) f_0(\underline{\theta}) d\underline{\theta}}{p_c(\overline{\mathcal{I}}^*_{0,t}|H_0, \underline{\theta}^*_0) \int_S p_c(\overline{\mathcal{I}}^*_{0,t}|H_1, \underline{\theta}) f_1(\underline{\theta}) d\underline{\theta}} \tag{4.35}$$

where

$$\underline{\theta}^*_i = \arg \max_{\underline{\theta}} p_c(\overline{\mathcal{I}}^*_{0,t}|H_i, \underline{\theta}) f_i(\underline{\theta}) \overline{\mathcal{I}}^*_{i,t} = \arg \max_{\mathcal{I}} p_t(\mathcal{I}|H_i, \underline{\theta}^*_i) \tag{4.36}$$

and $\underline{\theta}^*_0$ is the true nuisance parameters of the target.

The pseudo log likelihood ratio should be derived using the classifier noise model, but its resulting distribution is determined using the true distribution of the sample images. Therefore, probability of error $P(D_1|H_0)$ for this classifier can be calculated as

$$P(D_1|H_0) \sim \int_{\log \nu'}^{\infty} p(l|H_0) dl$$

where from (4.27)

$$p(l|H_0) = p(L'_c(\mathcal{I}) = l|p_t(\mathcal{I}|H_0, \underline{\theta}^*_0)) \tag{4.37}$$

and $L'_c(\mathcal{I})$ is the pseudo loglikelihood ratio used in (4.26) computed using the classifier pdfs p_c.

Next the Unbiased Integration Method is considered. The expectation in the threshold adjustment is the expectation of the likelihood given the true target distribution. Thus, the likelihood for which the expectation is computed is that given by the classifier noise model, but the true distribution is used as the underlying distribution with which the expectation is calculated.

$$\nu' = \nu \frac{E\left[p_c(\mathcal{I}|H_1, \underline{\theta}^*_1)|p_t(\mathcal{I}|H_0, \underline{\theta}^*_0)\right] \int_S E\left[p_c(\mathcal{I}|H_0, \underline{\theta})|p_t(\mathcal{I}|H_0, \underline{\theta}^*_0)\right] f_0(\underline{\theta}) d\underline{\theta}}{E\left[p_c(\mathcal{I}|H_0, \underline{\theta}^*_0)|p_t(\mathcal{I}|H_0, \underline{\theta}^*_0)\right] \int_S E\left[p_c(\mathcal{I}|H_1, \underline{\theta})|p_t(\mathcal{I}|H_0, \underline{\theta}^*_0)\right] f_1(\underline{\theta}) d\underline{\theta}} \tag{4.38}$$

As a result, the Unbiased Integration Method is capable of integrating the knowledge of the correct noise model into the calculation of the threshold adjustment, whereas the Noiseless Integration Method is not. Hence, as will be seen, the Unbiased Integration prediction tends to greatly outperform the Noiseless Integration method for this type of problem. The pseudo log likelihood ratio distribution is computed in the same way as for the Noiseless Integration Method.

Chapter 5

Implementation

In this chapter, we discuss our implementation of the derived Bayesian performance prediction methods in realistic scenarios with arbitrary noise and nuisance parameters.

5.1 Computing the Integrals

5.1.1 Method of Grenander

As the method of Grenander uses the Laplace approximation to evaluate the required integrals to find ν' (4.13), the second derivatives of the logarithms of the integrands need to be found at the peaks, as described in the derivation of the method. The required second derivative in (4.12) can be obtained using finite differencing the expression in (6.11) about the true pose for Target 0 and about the pose of closest approach for Target 1 using the method of [3]. The spacing of the points used to perform the finite differencing can be selected as desired, keeping in mind that the use of the second derivative in this method is based on the assumption that the function is close to a Gaussian pulse, and that rapid variations of the function about its overall trajectory are not important. For our experiment, 5 points evenly and symmetrically spaced around the maximum point were used, with the outermost points chosen such that they had values approximately 90% of the value of the maximum point. The pose of closest approach is the pose for which the likelihood of Target 1 given

Target 0 in its true pose is maximized.

5.1.2 NIM and UIM

In order to compute the threshold ν' for either the Noiseless Integration Method (NIM) or the Unbiased Integration Method (UIM), deterministic likelihood integrals found in (4.15) or (4.38) must be computed numerically. To do this, the integrands need to be evaluated for a set of discrete values of $\underline{\theta}$, chosen in such a way that the resulting numerical integral will have the desired accuracy. Each of these points requires the rendering of a noiseless LADAR image corresponding to selected target and pose. Since the logarithms of the curves to be integrated are asymptotically quadratic and thus have a clearly defined peak, we use simple optimization techniques that are able to avoid local maxima to determine the peak and the width of the portion of the integrand that has sufficient magnitude. Using this, we then determine the desired spacing of the sample points, typically using a specified number of points.

The integrand (4.15) required by Section 4.4 can be straightforwardly evaluated, since it simply involves calculating the likelihood of one noiseless image given that a different noiseless image is true, which is done by a direct evaluation of the pdf of the noise.

The expectations in the integrands (4.38) for the Unbiased Integration Method are more difficult to evaluate. In several cases of interest (see Section 3), the expectations can be calculated analytically, allowing the integrands to be calculated by a function evaluation. If an analytic expression is unavailable, it is possible to obtain the expectations via numerical integration. However, this is of course computationally expensive, and will likely result in the prediction algorithm not significantly outperforming Monte Carlo integration in terms of computational complexity. Thus, the Unbiased Integration Method should not be used when the required expectations cannot be analytically calculated and the method of Section 4.3 or the Method of Noiseless Integration should be used instead.

Once the integrands have been evaluated for a number of discrete values of $\underline{\theta}$, the required R-dimensional integrals ((4.22) or (4.9), where R is the dimension of $\underline{\theta}$) can be approximated by performing numerical integration. Since the integrands asymptotically become Gaussian functions [1],

we found it natural to use Gaussian interpolation and extrapolation to perform the numerical integration instead of the linear interpolation used by the ubiquitous trapezoidal rule, so as to keep the required integrand evaluations to a minimum. For portions of the tail for which Gaussian interpolation is impossible due to the positive second derivative of the logarithm, exponential interpolation was used instead. In our experience, these interpolation techniques are significantly more accurate than linear interpolation for the computation of the integral.

As this numerical integration is computationally identical to the likelihood integral in the Bayesian classifier, the computational complexity of this portion of the prediction is on the same order of magnitude of the performance of one Bayesian ATR operation.

5.1.3 Selection of Samples for Multi-Dimensional Numerical Integration

In order to perform the requisite multi-dimensional integration to obtain the threshold adjustment, it is necessary to determine the location and spacing of the needed samples. Since the integrands to compute the threshold adjustments asymptotically have a Gaussian shape, the integrands almost always have a general peaked shape with low tails, although with near symmetries in the target several peaks can develop. As a result, for efficient sampling, frequently only the neighborhood of the peaks needs to be sampled. The location and width of this neighborhood is not known a priori, with the exception of the location of the primary peak for the true target, which of course is located at the true pose based on our assumption of consistency of inference.

Thus, for this type of efficient sampling to be performed, an optimization-based technique to discover the location of these peaks and their size is required. Since the integrand is a likelihood or an expectation of a likelihood, we work in the log likelihood domain, so as to be able to find the peak even if the initial guess has a rather low magnitude. For our experiments, we used an algorithm based on a fusion of gradient descent and iteratively fitting a quadratic function to a samples that gradually converge on the peak. The latter addition was needed to avoid the large number of local maxima. Note that it is not critical for our methods to find the exact maximum at this stage, because the entire peak will be sampled once its region for integration is determined. We found the width

of the peak by using the a quadratic fit to get an initial estimate and then expanding the width until the samples indicated that the magnitude of the function was sufficiently small. Other methods applicable to the finding of the peak include MCMC based methods, such as that proposed in [37], which we do not discuss further.

Once the region over which to integrate is found, it remains to evenly sample over the region. In some cases, it may be possible to approximate the multidimensional integral associated with each peak as separable into the product of one dimensional integrals along lines through the overall peak of the integrand and parallel to the coordinate axes. This would allow the reduction in the number of samples by allowing the sampling along only a set of mutually perpendicular lines. This method for finding the regions over which to integrate was found to be reasonably robust in our experiments, and converged at a reasonable rate.

5.2 Computing the Pseudo LLR PDF

The remaining task in calculating the performance using the threshold adjusting methods is to derive the pdf of the pseudo log-likelihood ratio on the left side of (4.12). This pdf will then be used to compute the probability of error using the separately computed threshold ν'. It is possible to analytically derive this for the case of Gaussian noise and some other noise types. For most types of noise, however, it is not possible to obtain a closed form expression for the pdf. We present two possible approaches of approximating the required pdf, using the model that the noise for each of the N pixels are independently distributed. First, let there be N K-dimensional pixels in the image $\mathcal{I}$, each denoted as $\underline{s_k},\ l = 1, ..., N$.

The independence of the pixels allows the decomposition of the pseudo LLR into a sum of the

individual pixel LLRs.

$$\begin{aligned} L'(\mathcal{I}) &= (\log p(\mathcal{I}|H_1, \underline{\theta}_1^*) - \log p(\mathcal{I}|H_0, \underline{\theta}_0^*)) \\ &= \sum_{k=1}^{N} (\log p(\underline{s}_k|H_1, \underline{\theta}_1^*) - \log p(\underline{s}_k|H_0, \underline{\theta}_0^*)) \\ &= \sum_{k=1}^{N} L'_k(\underline{s}_k) \end{aligned} \tag{5.1}$$

One method relies on approximating the pdf as Gaussian. Since the pixel noises are independent, the L'_k are a set of N independent random variables. As a result, the Central Limit Theorem comes into effect for N sufficiently large, allowing the pdf to be approximated by a Gaussian distribution. Since the pixels are not IID, the additional assumption that the Lindeberg condition holds is required [8]. In this case, the mean and variance of the distribution can be found by adding the means and variances of the log-likelihood ratio $L'_k(\underline{s}_k)$ associated with each pixel, thus uniquely determining the pdf of the overall log-likelihood ratio $L'(\mathcal{I})$.

$$\begin{aligned} E\left[L'(\mathcal{I})\right] &= \sum_{j=1}^{N} E\left[L'_k(\underline{s}_k)\right] \\ \mathrm{Var}\left[L'(\mathcal{I})\right] &= \sum_{k=1}^{N} \mathrm{Var}\left[L'_k(\underline{s}_k)\right] \end{aligned} \tag{5.2}$$

This gives the simple expression for the probability of error

$$P_e \approx Q\left(\frac{\nu' - E[L'(\mathcal{I})]}{\sqrt{\mathrm{Var}[L'(\mathcal{I})]}}\right) \tag{5.3}$$

where as before ν' is the threshold computed using one of the prediction methods.

If this method cannot be applied, the second proposed method is to calculate the pdf numerically. The pdfs of the log-likelihood ratios for each pixel can be found using the noise pdf and the inversion method. Since the pixel noises are independent, the overall pdf can then be found by numerically convolving each of the pixel level pdfs. Difficulties with this method are that the pixel pdfs tend to contain impulses, which may require relatively large numbers of points to describe accurately. On the other hand, the convolution operation can be done quite efficiently by convolving

the pdfs in a hierarchical fashion and downsampling as the functions become oversampled as they become smoother and smoother.

The last method is to compute the probability of error directly using a Monte Carlo simulation. Since the simulation involves no nuisance parameter uncertainty, it can be done quite efficiently and robustly. Moreover, it is clear that this simulation is much more efficient than is the Monte Carlo simulation of the entire ATR system, since no Bayesian integrals are required in the simulation.

5.3 Approximating the Accuracy of the Prediction

For almost any prediction of any type of performance, it is critical to know how accurate the prediction is. For the problem under consideration, a situation may arise where high prediction accuracy is required, thus necessitating an approximation of the accuracy of the chosen prediction method.

Since arbitrary accuracy can be achieved using the multi-modal approximation technique discussed in Section 4.7.1 for a sufficiently large R_i, we propose the following method of approximating the accuracy of a prediction of probability of error. First, compute the prediction using the method whose accuracy is being tested. Next, compute the prediction with using the multi-modal technique with small R_i and appropriately selected random samples. Increase R_i, reselect the points and recompute the prediction. Continue increasing until the predicted values begin to converge, with a threshold determined by the required accuracy of the error calculation. The probability of error to which this series has converged is then used as the true performance, and the difference between this and the prediction using the method under test gives the estimated prediction error. Of course, to ensure that a false convergence is not achieved, it is critical that the locations of the approximating random samples $\underline{\theta}_i^{(w)}$ for the multi-modal prediction are always chosen so that they are approximately evenly spaced throughout the portion of the region S for which the relevant integrand is large enough to contribute meaningfully to the integral.

Note that since the accuracy of a prediction should vary smoothly with the variation of the ATR scenario parameters, it is only necessary to determine the more accurate prediction for a few parameter values in order to estimate the error over the relevant range of parameter values for which

the prediction will be computed.

Chapter 6

Application to LADAR Noise Models

This chapter applies the prediction methods of Chapter 4 to the noise models of Chapter 2.6. To calculate the prediction using either method, the threshold ν' (or the weights ν'_{ir}) and the pseudo LLR pdf $p(l|H_0)$ of L' must be computed. Once these are found, the probabilities of errors can be found using (4.27) or (4.34).

For each method, the ν' or ν'_{ir} are computed by the integration of certain functions. Hence, in order to apply the prediction methods to each noise model, we must derive the integrands. The integrands (excluding the priors f_i, which are specified directly) for the method of Grenander and the NIM (4.10),(4.15) are given by

$$p(\overline{\mathcal{I}}_0^*|H_i, \underline{\theta}) \tag{6.1}$$

for all i and $\underline{\theta}$. As this is merely the likelihood function for each noise model, it can be computed using the defining likelihood function and the appropriate rendered model image.

For the UIM, a different quantity is needed for the integrand (4.38). This is given by the expectation

$$E\left[p(\mathcal{I}|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*\right] \tag{6.2}$$

Moreover, all the models presented in Chapter 2.6 have independent pixel distributions. Hence, we have that

$$E\left[p(\mathcal{I}|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*\right] = \prod_{k=1}^{N} E\left[p(\underline{s}_k|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*\right] \tag{6.3}$$

In this chapter, we thus derive the quantity

$$E\left[p(\underline{s}_k|H_i,\underline{\theta})|H_0,\underline{\theta}_0^*\right] \tag{6.4}$$

for each noise model to complete the prediction. This is given by

$$E\left[p(\underline{s}_k|H_i,\underline{\theta})|H_0,\underline{\theta}_0^*\right] = \int_{-\infty}^{\infty} p(\underline{s}_k|H_i,\underline{\theta})p(\underline{s}_k|H_0,\underline{\theta}_0^*)d\underline{s}_k \tag{6.5}$$

where the true underlying distribution of $\underline{s}_k$ is used.

Moving to the pdf of the pseudo loglikelihood ratio (4.26),

$$L' = \log\frac{p(\mathcal{I}|H_1,\underline{\theta}_1^*)}{p(\mathcal{I}|H_0,\underline{\theta}_0^*)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \log\nu'$$

For some noise models it is possible to analytically compute its pdf. Hence, in this chapter, we also derive the pdf of L' where possible.

6.1 Noise Model Components

In this section, we apply the Bayesian performance prediction methods discussed up to this point to several simple noise models. These results will be used as building blocks for the prediction using the specifically LADAR noise models discussed in the next section. First, we will predict the performance of multi-dimensional arbitrary Gaussian noise that is independent from pixel to pixel. Secondly, we will derive prediction results for general mixture PDF models. Finally, we will use a general random pixel deletion model.

6.1.1 Gaussian Noise

Since Gaussian noise is part of many of the LADAR noise models discussed in this thesis, it is necessary to be able to predict performance using the Gaussian noise model. These results will then be applied to the several LADAR noise models that utilize Gaussian noise.

Consider the case where the image $\mathcal{I}$ has K dimensional pixels with a pixel noise model that is additive independent Gaussian noise with $(K \times K)$ covariance $\mathbf{\Sigma}_{k,i}$ under hypothesis H_i for each pixel $k = 1, ..., N$. This covariance can vary from pixel to pixel and as a function of hypothesis and nuisance parameter values, that is $\mathbf{\Sigma}_{k,i} = \mathbf{\Sigma}_{k,i}(\underline{\theta})$. Then, the likelihood function is a product of K-dimensional Gaussian pdfs. Thus, assuming hypothesis H_i and nuisance parameters $\underline{\theta}$, the pixel observations $\underline{s}_k$ are distributed as $\mathcal{N}(\overline{\underline{s}}_{k,i}(\underline{\theta}), \mathbf{\Sigma}_{k,i})$, giving an overall image likelihood of [9]

$$p(\mathcal{I}|H_i, \underline{\theta}) = \frac{1}{\left((2\pi)^K |\mathbf{\Sigma}_{k,i}(\underline{\theta})|\right)^{N/2}} \times \exp\left\{ -\frac{1}{2} \sum_{k=1}^{N} (\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta}))^T \mathbf{\Sigma}_{k,i}(\underline{\theta})^{-1} (\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta})) \right\} \tag{6.6}$$

UIM Integrand Calculation

For UIM, in order to compute ν' in (4.38) it is necessary to find an expression for the expectation $E[p(\mathcal{I}|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*]$. Using the expression in (6.6), we have

$$\begin{aligned} &E[p(\underline{s}_k|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*] = \\ &\frac{1}{(2\pi)^K \sqrt{|\mathbf{\Sigma}_{k,i}||\mathbf{\Sigma}_{k,i}|}} \int_{-\infty}^{\infty} \left[\exp\left\{ -\frac{1}{2} (\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta}))^T \mathbf{\Sigma}_{k,i}^{-1} (\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta})) \right\} \right. \\ &\left. \cdot \exp\left\{ -\frac{1}{2} (\underline{s}_k - \overline{\underline{s}}_{k,0}^*)^T \mathbf{\Sigma}_{k,0}^{-1} (\underline{s}_k - \overline{\underline{s}}_{k,0}^*) \right\} \right] d\underline{s}_k \\ &= \frac{1}{(2\pi)^K \sqrt{|\mathbf{\Sigma}_{k,i}||\mathbf{\Sigma}_{k,0}|}} \int_{-\infty}^{\infty} \exp\left\{ -\frac{1}{2} \left((\underline{s}_k - \mathbf{A}_k^{-1} \underline{\mu}_k)^T \mathbf{A}_k (\underline{s}_k - \mathbf{A}_k^{-1} \underline{\mu}_k) + \xi_k \right) \right\} d\underline{s}_k \end{aligned} \tag{6.7}$$

where

$$\begin{aligned}\mathbf{A}_k &= \mathbf{\Sigma}_{k,i}^{-1} + \mathbf{\Sigma}_{k,0}^{-1} \\ \underline{\mu}_k &= \mathbf{\Sigma}_{k,i}^{-1}\overline{\underline{s}}_{k,i}(\underline{\theta}) + \mathbf{\Sigma}_{k,0}^{-1}\overline{\underline{s}}_{k,0}^{*} \\ \xi_k &= -\underline{\mu}_k^T \mathbf{A}_k^{-1}\underline{\mu}_k + \overline{\underline{s}}_{k,0}^{*T}\mathbf{\Sigma}_{k,0}^{-1}\overline{\underline{s}}_{k,0}^{*} + \overline{\underline{s}}_{k,i}(\underline{\theta})^T\mathbf{\Sigma}_{k,i}^{-1}\overline{\underline{s}}_{k,i}(\underline{\theta})\end{aligned} \tag{6.8}$$

since $\mathbf{\Sigma}_{k,i}$ is symmetric by definition. Using the standard result for the integral of a Gaussian curve, this results in

$$E[p(\mathcal{I}|H_i,\underline{\theta})|H_0,\underline{\theta}_0^*] = \frac{\exp\left\{\sum_{k=1}^{N}\xi_k\right\}}{(2\pi)^{KN/2}\sqrt{\prod_{k=1}^{N}|\mathbf{\Sigma}_{k,i}||\mathbf{\Sigma}_{k,0}||\mathbf{A}_k|}} \tag{6.9}$$

As can be seen from the LADAR noise models discussed so far, a useful special case of the Gaussian noise model is the case where $\mathbf{\Sigma}_{k,i} = \mathbf{\Sigma}_{k,0}, \forall k$. When this is the case, the expression in (6.9) simplifies to

$$E[p(\underline{s}_k|H_i,\underline{\theta})|H_0,\underline{\theta}_0^*] = \frac{1}{(2(2\pi)^K|\mathbf{\Sigma}_k|)^{N/2}}\exp\left\{-\frac{1}{4}\left(\overline{\underline{s}}_{k,0}^{*} - \overline{\underline{s}}_{k,i}(\underline{\theta})\right)^T\mathbf{\Sigma}_k^{-1}\left(\overline{\underline{s}}_{k,0}^{*} - \overline{\underline{s}}_{k,i}(\underline{\theta})\right)\right\} \tag{6.10}$$

Pseudo LLR Distribution for Equal Covariance

Once the modified threshold ν is obtained, it remains to find the pdf of the pseudo LLR $L'(\mathcal{I}) = -2|\mathbf{\Sigma}| = E_1(\mathcal{I},\underline{\theta}_1^*,\mathbf{\Sigma}_{k,1}) - E_0(\mathcal{I},\underline{\theta}_0^*,\mathbf{\Sigma}_{k,0})$.

In order that the L' will be Gaussian distributed, we consider only the case for which $\mathbf{\Sigma}_{k,1} = \mathbf{\Sigma}_{\mathbf{k,0}}, \forall j$. This case will be sufficient for the noise models that will be considered. Thus, we abbreviate the pixel covariance as merely $\mathbf{\Sigma}_k = \mathbf{\Sigma}_{k,i}, i = \{0,1\}$. This gives a log-likelihood (6.6)

of

$$E_i(\mathcal{I}, \underline{\theta}, \mathbf{\Sigma}_k) = -2\log \frac{1}{\sqrt{\left((2\pi)^{KN}\prod_{k=1}^{N}|\mathbf{\Sigma}_k|\right)}} + \sum_{j=1}^{N}(\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta}))^T \mathbf{\Sigma}_k^{-1}(\underline{s}_k - \overline{\underline{s}}_{k,i}(\underline{\theta})) \tag{6.11}$$

Thus, the log likelihood ratio for H_1 at pose $\underline{\theta}_1$ and H_0 at pose $\underline{\theta}_0$ is given by

$$\begin{aligned} L'(\mathcal{I}, \underline{\theta}_0, \underline{\theta}_1) &= \log \frac{p(I|H_1, \underline{\theta}_1)}{p(I|H_0, \underline{\theta}_0)} \\ &= -\frac{1}{2}\sum_{j=1}^{N}(\underline{s}_k - \overline{\underline{s}}_{k,1}(\underline{\theta}))^T \mathbf{\Sigma}_k^{-1}(\underline{s}_k - \overline{\underline{s}}_{k,1}(\underline{\theta})) \\ &\quad + \frac{1}{2}\sum_{j=1}^{N}(\underline{s}_k - \overline{\underline{s}}_{k,0}(\underline{\theta}))^T \mathbf{\Sigma}_k^{-1}(\underline{s}_k - \overline{\underline{s}}_{k,0}(\underline{\theta})) \\ &= -\sum_{j=1}^{N}(\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1))^T \mathbf{\Sigma}_k^{-1}\underline{s}_k \\ &\quad + \frac{1}{2}\sum_{j=1}^{N}\left(\overline{\underline{s}}_{k,0}(\underline{\theta}_0)^T \mathbf{\Sigma}_k^{-1}\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1)^T \mathbf{\Sigma}_k^{-1}\overline{\underline{s}}_{k,1}(\underline{\theta}_1)\right) \end{aligned} \tag{6.12}$$

Since the $\underline{s}_k$ are Gaussian, $L'(\mathcal{I})$ is also Gaussian with mean and variance

$$\begin{aligned} E[L'(\mathcal{I}, \underline{\theta}_0, \underline{\theta}_1)] &= -\sum_{j=1}^{N}(\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1))\mathbf{\Sigma}_k^{-1}\overline{\underline{s}}_0^{*(j)} \\ &\quad \frac{1}{2}\sum_{j=1}^{N}\left(\overline{\underline{s}}_{k,0}(\underline{\theta}_0)^T \mathbf{\Sigma}_k^{-1}\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1)^T \mathbf{\Sigma}_k^{-1}\overline{\underline{s}}_{k,1}(\underline{\theta}_1)\right) \\ \mathrm{Var}[LLR(\mathcal{I}, \underline{\theta}_0, \underline{\theta}_1)] &= \sum_{j=1}^{N}(\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1))^T \mathbf{\Sigma}_k^{-1}(\overline{\underline{s}}_{k,0}(\underline{\theta}_0) - \overline{\underline{s}}_{k,1}(\underline{\theta}_1)) \end{aligned} \tag{6.13}$$

6.1.2 Mixture PDF

Consider the case of noise pdfs composed of mixtures of other pdfs. This will allow the use of more complex noise models by decomposing the noise pdfs into mixtures of simple distributions, for which the required moments are already derived. This model will be useful later for LADAR

noise models incorporating more accurate random cross-range noise models.

Assume that the noise for each pixel is independent. Suppose that the pdf $p\left(\underline{s}_k|H_i,\underline{\theta}\right)$ of the kth pixel given H_i and nuisance parameters $\underline{\theta}$ is a mixture of R component pdfs $p_r(\underline{s}_k|H_i,\underline{\theta})$, $r=1,...,R$. To ensure generality, the number of components R is a function of i,j, and $\underline{\theta}$, that is, $R=R(i,k,\underline{\theta})$, with the functional dependencies left out for simplicity of notation. The weights associated with each pdf component are denoted as w_r, $r=1,...,R$. Again, the w_r are functions of i,k, and $\underline{\theta}$, or $w_r=w_r(i,k,\underline{\theta})$. By total probability, it is necessary that $\sum_{r=1}^{R} w_r=1$.

This gives a pixel pdf of

$$p\left(\underline{s}_k|H_i,\underline{\theta}\right)=\sum_{r=1}^{R} w_r p_r(\underline{s}_k|H_i,\underline{\theta}) \tag{6.14}$$

The complete image pdf is the product of the individual pixel pdfs by independence.

$$p\left(\mathcal{I}|H_i,\underline{\theta}\right)=\prod_{k=1}^{N} p\left(\underline{s}_k|H_i,\underline{\theta}\right) \tag{6.15}$$

The UIM requires the evaluation of the expectation of this likelihood function for each pixel given the true distribution $(H_0,\underline{\theta}^*)$ of the same form but possibly different parameters. This is given by

$$E\left[p\left(\underline{s}_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}_0^*\right]=\sum_{r_0=1}^{R_0}\sum_{r_i=1}^{R_1} w_{0,r_0}w_{i,r_i}E\left[p_{r_i}(\underline{s}_k|H_i,\underline{\theta})|p_{r_0}(\underline{s}_k|H_0,\underline{\theta}_0^*)\right] \tag{6.16}$$

where p_{r_i} is the r_ith component of the mixture distribution under H_i and appropriate nuisance parameters, R_i denotes the number of components, and w_{i,r_i} denotes the r_ith corresponding weights. This result follows from the linearity of the expectation operator.

Hence, if the required expectation is available for a certain distribution, then the expectation can be calculated for any distribution formed using mixtures of that distribution. This fact is useful since highly complex noise pdfs can be approximated by mixtures of simpler distributions.

6.1.3 Random Deletions

In LADAR imagery, anomalous pixels (described in Chapter 2.6) are a significant factor. These pixels are pixels that occur randomly throughout the image, and typically have pdfs that contain little to no information about the actual target. As a result, some have used a reasonable approximation that these pixels are to be treated as deleted pixels [7]. In essence, this is an approximation that assumes that virtually all anomalous pixels can be identified and removed by the classification algorithm, and the pdf of the anomalous pixels provides no, or at least negligible information about the target. Given that anomalous pixels are typically distributed approximately uniformly over the entire range interval, these assumptions should be accurate in most cases [7]. This approximation is attractive from a prediction standpoint because it simplifies the required calculations considerably for most types of false alarms, especially when computing the expectations required for the Method of Unbiased Integration. In Appendix B, we demonstrate, as an example, that the Gaussian noise with random deletions model is asymptotically equivalent to the coherent LADAR model. In a later section, we will also provide experimental verification of the accuracy of using the deletion model as an approximation for a uniformly distributed anomaly model.

This model should also be useful in the modeling of obscuration effects. If an object between the target and the sensor is obscuring part of the target, it is sometimes the case that the obscuring object is known to not be part of the target due to its distance from the target, as determined by the use of a detection algorithm [38]. If this is not the case, then the effect of the obscuring object would be highly dependent on its shape, and thus should be included in the CAD models for the rendering of the noiseless imagery determining the target image distributions. An example scenario with obscuring objects that are equivalent to deletions would be one for which a high tree canopy is between the target and the sensor [38]. Leaves or branches in the canopy obscure many of the LADAR pixels, but not all [38]. Since the leaves are much higher than the target, there is virtually no chance of mistakenly associating them with the target due to the extremely low likelihood that target pixels would have such large errors, given reasonable sensor noise levels [38]. If this is indeed so, it is clear that the obscured pixels can be modeled as deletions with minimal error [8]. Moreover, the true distribution of the obscuring objects is probably not known a priori, thus further suggesting

the use of a deletion model. Thus, random obscuration can be modeled using this method of random deletions with probabilities of deletions that may vary across the image, depending on the desired model.

Assume as before that the noise for each pixel is independent, with the kth pixel having a non-anomalous pdf under H_i and nuisance parameters $\underline{\theta}$ of $p_{na}\left(\underline{s}_k|H_i,\underline{\theta}\right)$ over the set $\mathcal{A}$ and a probability of deletion $\alpha_{k,i}(\underline{\theta})$. Thus, the noise model is

$$p\left(\underline{s}_k|H_i,\underline{\theta}\right)=\begin{cases}\left(1-\alpha_{k,i}(\underline{\theta})\right)p_{na}\left(\underline{s}_k|H_i,\underline{\theta}\right) & \underline{s}_k\in\mathcal{A}\\ \frac{\alpha_{k,i}(\underline{\theta})}{|\mathcal{B}|} & \underline{s}_k\in\mathcal{B}\end{cases} \tag{6.17}$$

The expectation of this pdf given the true hypothesis H_0 at the true nuisance parameters $\underline{\theta}^*$ is required for the Method of Unbiased Integration. This is given by

$$E\left[p\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}^*\right]=\left(1-\alpha_{k,0}(\underline{\theta}^*)\right)\left(1-\alpha_{k,i}(\underline{\theta})\right)E\left[p_{na}\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}^*\right] \tag{6.18}$$

where we let $|\mathcal{B}|$ approach infinity with no change to the model. This choice is based on the fact that the volume of $\mathcal{B}$ has no effect on the actual likelihood ratio, and has the added benefit that it prevents deleted pixels from penalizing the likelihood ratio. This incorrect penalty results because the expectation, being additive, would not cancel out the anomalous probabilities as happens with the true likelihood ratio. The lack of cancelation results from the unbiased nature of the approximation for ν. If the anomalous likelihood is made overly large by collapsing $\mathcal{B}$, this in effect shifts the impulse associated with the anomalies in the pdf of the likelihood far in the positive direction, thus artificially increasing the expectation of the likelihood and destroying the accuracy of the approximation of the remainder of the pdf. This is important because the remainder of the pdf is the only portion that affects the overall likelihood ratio.

The pixel likelihood ratio between hypotheses i_0 and i_1 at nuisance parameter values of $\underline{\theta}_0$ and

$\underline{\theta}_1$ respectively is given by

$$L'_k(s_k|\underline{\theta}_0,\underline{\theta}_1) = \log\frac{p\left(\underline{s}_k|H_{i_1},\underline{\theta}_1\right)}{p\left(\underline{s}_k|H_{i_0},\underline{\theta}_0\right)} \tag{6.19}$$
$$= \begin{cases} \log\frac{\left(1-\alpha_{k,i_1}(\underline{\theta}_1)\right)}{\left(1-\alpha_{k,i_0}(\underline{\theta}_0)\right)}\frac{p_{na}\left(\underline{s}_k|H_{i_1},\underline{\theta}_1\right)}{p_{na}\left(\underline{s}_k|H_{i_0},\underline{\theta}_0\right)} & \text{Pixel not deleted} \\ \log\frac{\alpha_{k,i_1}(\underline{\theta}_1)}{\alpha_{k,i_0}(\underline{\theta}_0)} & \text{Pixel deleted} \end{cases}$$

and by independence the complete likelihood ratio is

$$L'(I|\underline{\theta}_0,\underline{\theta}_1) = \sum_{k=1}^{N} L'_k(\underline{s}_k|\underline{\theta}_0,\underline{\theta}_1) \tag{6.20}$$

To do performance prediction, the pdf of L' is needed. First, the pdf of L'_k needs to be obtained. Let $p^{(k)}_{L,na}(\ell)$ be the pdf of the non-anomalous single pixel log-likelihood ratio

$$L'_{k,na}(\underline{s}_k|\underline{\theta}_0,\underline{\theta}_1) = \log\frac{p_{na}\left(\underline{s}_k|H_{i_1},\underline{\theta}_1\right)}{p_{na}\left(\underline{s}_k|H_{i_0},\underline{\theta}_0\right)} \tag{6.21}$$

The distribution of L' is of course its distribution given the image $\mathcal{I}$ is of the true hypothesis H_0 with the true nuisance parameters $\underline{\theta}^*$. This pdf is typically easier to obtain than the pdf of the log-likelihood ratio for the complete noise including false alarms since the relevant noise pdfs are simpler for most noise models. Then, the pdf of $L'_k(s_k|\underline{\theta}_0,\underline{\theta}_1)$ is given by

$$p^{(k)}_L(\ell) = \left(1-\alpha_{k,0}(\underline{\theta}^*)\right)p^{(k)}_{L,na}\left(\ell - \log\left(\frac{1-\alpha_{k,i_1}(\underline{\theta}_1)}{1-\alpha_{k,i_0}(\underline{\theta}_0)}\right)\right) \tag{6.22}$$
$$+\,\alpha_{k,0}(\underline{\theta}^*)\delta\left(\ell - \log\left(\frac{\alpha_{k,i_1}(\underline{\theta}_1)}{\alpha_{k,i_0}(\underline{\theta}_0)}\right)\right)$$

and by independence the pdf of the complete LLR can be obtained using

$$p_L(\ell) = p^{(1)}_L(\ell) * \ldots * p^{(N)}_L(\ell) \tag{6.23}$$

where $*$ denotes convolution. This result will likely reduce the computational burden in computing the required LLR pdf for the case that the pdf of the non-anomalous LLR can be analytically calculated, e.g. for the Gaussian noise with anomalous pixels model. This is because it provides an approximation for which the pdfs of the per pixel LLRs can be analytically calculated, eliminating

the need for the use of numerical methods to determine the pdfs based on the pdfs of the underlying pixel distributions.

Alternatively, if the number of pixels in the image is sufficiently high such that the Central Limit Theorem applies, the distribution of the LLR can be approximated as Gaussian. We thus only need to specify its mean and variance. Let $E_{na}(k)$ be the expectation of $L'_{j,na}(\underline{s}_k|\underline{\theta}_0,\underline{\theta}_1)$ and $\sigma^2_{na}(k)$ its variance. From (6.22), we have that

$$E_L(k) = (1-\alpha_{k,0}(\underline{\theta}^*))\,E_{na}(k) + \alpha_{k,0}(\underline{\theta}^*)\log\left(\frac{\alpha_{k,i_1}(\underline{\theta}_1)}{\alpha_{k,i_0}(\underline{\theta}_0)}\right) \tag{6.24}$$

$$\begin{aligned}\sigma^2_L(k) &= (1-\alpha_{k,0}(\underline{\theta}^*))\,\sigma^2_{na}(j)\\ &\quad + \left((1-\alpha_{k,0}(\underline{\theta}^*)) - (1-\alpha_{k,0}(\underline{\theta}^*))^2\right)E^2_{na}(j)\left(\alpha_{k,0}(\underline{\theta}^*) - \alpha_{k,0}(\underline{\theta}^*)^2\right)\left(\log\left(\frac{\alpha_{k,i_1}(\underline{\theta}_1)}{\alpha_{k,i_0}(\underline{\theta}_0)}\right)\right)^2\\ &\quad - 2\,(1-\alpha_{k,0}(\underline{\theta}^*))\,E_{na}(j)\alpha_{k,0}(\underline{\theta}^*)\log\left(\frac{\alpha_{k,i_1}(\underline{\theta}_1)}{\alpha_{k,i_0}(\underline{\theta}_0)}\right)\end{aligned}$$

where $E_L(k)$ and $\sigma_L(k)$ are the complete LLR pixelwise expectation and standard deviation. Using these and the independence of the pixels, we of course have that the overall expectation and variance are

$$E_L = \sum_k E_L(k)$$
$$\sigma^2_L = \sum_k \sigma^2_L(k)$$

which thus determine the Gaussian pdf approximation to the LLR given that of the nonanomalous LLR.

Using this, we have that the pseudo-loglikelihood based prediction of the performance is given by (4.26)

$$P(D_1|H_0) \approx Q\left(\frac{\nu' - E[L']}{\sqrt{\mathrm{Var}[L']}}\right) \tag{6.25}$$

where ν' is determined by one of the three prediction methods.

Performance with Constant Probability of Anomaly

Suppose now that the α are constant over $i, k, \underline{\theta}$, which is frequently approximately the case. This gives from (6.24) that

$$E_L = (1-\alpha)E_{na} \tag{6.26}$$
$$\sigma_L^2 = (1-\alpha)\,\sigma_{na}^2$$

where

$$E_{na} = \sum_k E_{na}(k)$$
$$\sigma_{na}^2 = \sum_k \sigma_{na}^2(k)$$

are the moments of the pseudo LLR when the probability of anomaly is zero. The threshold ν' is computed using the expectation, now given by

$$E\left[p\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}^*\right] = (1-\alpha)^2\,E\left[p_{na}\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}^*\right] \tag{6.27}$$

Since ν' depends on the ratio of integrals of this function, and the only effect that varying the probability of anomaly has is to scale the integrand evenly, the variation of the probability of anomaly does not affect ν', that is $\nu'(\alpha) = \nu'$.

The probability of error is then given by (4.27)

$$P(D_1|H_0) \approx Q\left(\frac{\nu' - (1-\alpha)E_{na}}{\sqrt{(1-\alpha)\,\sigma_{na}^2}}\right) \tag{6.28}$$

where ν' is a constant. The simplicity of this expression is one of the primary reasons for using the Random Deletion anomaly model.

6.2 LADAR Noise Models

In this section, we apply the performance prediction methods discussed in this paper to several more realistic LADAR noise models.

We discuss two common LADAR noise models here: the typical coherent LADAR noise model, the Geiger-mode noise model, a simpler Gaussian noise with random deletions model, and a more complex noise model incorporating finite-beamwidth effects.

6.2.1 Gaussian Noise with False Alarms - Coherent LADAR

From Chapter 3, the pdf of the kth pixel is given by

$$p\left(s_k|H_i\right) = \tag{6.29}$$
$$\begin{cases} \frac{\alpha_{k,i}}{b-a} + \frac{1-\alpha_{k,i}}{\left(Q\left(\frac{a-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right)-Q\left(\frac{b-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right)\right)\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(s_k-\overline{s}_{k,i}(\underline{\theta}))^2}{2\sigma_k^2}\right\} & a < s_k < b \\ 0 \quad \text{otherwise} \end{cases}$$

For the Unbiased Integration Method, the expectation of this pdf given the true hypothesis H_0 at the true nuisance parameters $\underline{\theta}^*$ is required. The required expectation is that of the individual hypothesized pixel likelihood given the true distribution of the pixel. Combining the results for mixture pdfs (6.16) and the results for Gaussian noise (6.10), we have

$$\begin{aligned} & E\left[p\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}^*\right] \qquad (6.30) \\ &= \int_{-\infty}^{\infty} p\left(s_k|H_i,\underline{\theta}\right) p\left(s_k|H_0,\underline{\theta}^*\right) ds_k \\ &= \frac{\alpha_{k,i}(1-\alpha_{k,0}) + \alpha_{k,0}(1-\alpha_{k,i}) - \alpha_{k,i}\alpha_{k,0}}{(b-a)} \\ &\quad + \frac{(1-\alpha_{k,i})(1-\alpha_{k,0})\left(Q\left(\frac{2a-\overline{s}_{k,i}(\underline{\theta})-\overline{s}^*_{k,0}}{\sigma_k}\right) - Q\left(\frac{2b-\overline{s}_{k,i}(\underline{\theta})-\overline{s}^*_{k,0}}{\sigma_k}\right)\right)}{\left(Q\left(\frac{a-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right) - Q\left(\frac{b-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right)\right)\left(Q\left(\frac{a-\overline{s}^*_{k,0}}{\sigma_k}\right) - Q\left(\frac{b-\overline{s}^*_{k,0}}{\sigma_k}\right)\right)\sqrt{4\pi\sigma_k^2}} \\ &\quad \times \exp\left\{-\frac{(\overline{s}^*_{k,0} - \overline{s}_{k,i}(\underline{\theta}))^2}{4\sigma_k^2}\right\} \end{aligned}$$

where the first term is found by

$$\int_a^b \frac{1}{(b-a)^2} ds_k = \frac{1}{b-a} \tag{6.31}$$

for the uniform component expectation given the true uniform and

$$\int_a^b \frac{1}{b-a} \frac{1}{\left(Q\left(\frac{a-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right) - Q\left(\frac{b-\overline{s}_{k,i}(\underline{\theta})}{\sigma_k}\right)\right)\sqrt{2\pi\sigma_k^2}} \exp\left\{-\frac{(s_k - \overline{s}_{k,i}(\underline{\theta}))^2}{2\sigma_k^2}\right\} ds_k = \frac{1}{b-a} \tag{6.32}$$

for the expectation of the uniform component given the Gaussian component and vice versa.

6.2.2 Gaussian Noise with Geiger mode anomalies

In Section 3.2.2 the pixel noise pdf for Geiger mode LADAR was expressed as

$$p\left(s_k|H_i\right) = \alpha_{k,i} p_a(s_k) + \rho_{k,i} p_g(s_k|H_i, \underline{\theta}) + P_{del}^{(k,i)} \delta(s_k - b), \quad a < s_k < b$$

where

$$p_g(s_k|H_i, \underline{\theta}) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(s_k - \overline{s}_{k,i}(\underline{\theta}))^2}{2\sigma_k^2}} \quad a < s_k < b$$

$$p_a(s_k|H_i, \underline{\theta}) = \begin{cases} \frac{\alpha_{k,i}^{(1)}}{\alpha_{k,i}} \gamma_1^{(k,i)} \exp(-\beta s_k) & a < s_k < \overline{s}_{k,i}(\underline{\theta}) \\ \frac{\alpha_{k,i}^{(2)}}{\alpha_{k,i}} \gamma_2^{(k,i)} \exp(-\beta s_k) & \overline{s}_{k,i}(\underline{\theta}) < s_k < b \end{cases}$$

with the parameters defined in Section 3.2.2.

To implement the Unbiased Integration Method we need the expectation

$$E\left[p\left(s_k|H_i, \underline{\theta}\right)|H_0, \underline{\theta}_0^*\right] \tag{6.33}$$

The pixel pdf is a mixture of four distributions. We thus make use of the Mixture pdf results in Section 6.1.2 in (6.16). Note that the component weighted by P_{del} is a deletion, and is treated

accordingly. Using this method, we obtain the result that

$$E[p\left(s_k|H_i,\underline{\theta}\right)|H_0,\underline{\theta}_0^*] = \alpha_{k,i}\alpha_{k,0}E_{ai,a0} + \alpha_{k,0}\rho_{k,i}E_{a0,ni} + \alpha_{k,i}\rho_{k,0}E_{ai,n0} + \rho_{k,0}\rho_{k,i}E_{n0,ni} \tag{6.34}$$

where the deletions do not factor into the expectation as noted in Section 6.1.3 and where from (6.10)

$$E_{n0,ni} = E[p_g(s_k|H_i,\underline{\theta})|p_g(s_k|H_0,\underline{\theta}_0^*)] = \frac{1}{\sqrt{2(2\pi)\sigma_k^2}}\exp\left\{-\frac{\left(\overline{s}_{k,0}^* - \overline{s}_{k,i}(\underline{\theta})\right)^2}{4\sigma_k^2}\right\} \tag{6.35}$$

Also, it can be shown

$$\begin{aligned} E_{ai,a0} = E[p_a(s_k|H_i,\underline{\theta})|p_a(s_k|H_0,\underline{\theta}_0^*)] &= \frac{\alpha_{k,i}^{(1)}\alpha_{k,0}^{(1)}}{\alpha_{k,i}\alpha_{k,0}}\zeta(a,\kappa_1,1,1) \\ &+ \frac{\alpha_{k,i}^{(j_1)}\alpha_{k,0}^{(j_2)}}{\alpha_{k,i}\alpha_{k,0}}\zeta(\kappa_1,\kappa_2,j_1,j_2) \\ &+ \frac{\alpha_{k,i}^{(2)}\alpha_{k,0}^{(2)}}{\alpha_{k,i}\alpha_{k,0}}\zeta(\kappa_2,b,2,2) \end{aligned} \tag{6.36}$$

where if $\underline{s}_{k,0}(\underline{\theta}) > \overline{s}_{k,i}(\underline{\theta})$, $j_1 = 2, j_2 = 1$ and otherwise $j_1 = 1, j_2 = 2$, $\kappa_1 = \min(\overline{s}_{k,0}(\underline{\theta}), \overline{s}_{k,i}(\underline{\theta}))$, $\kappa_2 = \max(\overline{s}_{k,0}(\underline{\theta}), \overline{s}_{k,i}(\underline{\theta}))$, and

$$\zeta(x,y,j,\ell) = \frac{\gamma_j^{(k,i)}\gamma_\ell^{(k,0)}}{2\beta}(e^{-2\beta x} - e^{-2\beta y}) \tag{6.37}$$

Finally,

$$E_{a_j,n_\ell} = E[p_a(s_k|H_j,\underline{\theta}_j)|p_g(s_k|H_\ell,\underline{\theta}_\ell)] = E[p_g(s_k|H_\ell,\underline{\theta}_\ell)|p_a(s_k|H_j,\underline{\theta}_j)] \tag{6.38}$$

$$\begin{aligned}
&= \frac{\alpha_{k,i}^{(1)}}{\alpha_{k,i}} \\
&\times \left(\gamma_1^{(k,i)} e^{\beta(\beta\sigma^2/2 - \overline{s}_{k,\ell}(\underline{\theta}))} \left(Q\left(\frac{a - \overline{s}_{k,\ell}(\underline{\theta}) + \beta\sigma^2}{\sigma}\right) - Q\left(\frac{\overline{s}_{k,j}(\underline{\theta}) - \overline{s}_{k,\ell}(\underline{\theta}) + \beta\sigma^2}{\sigma}\right)\right)\right) \\
&+ \frac{\alpha_{k,i}^{(2)}}{\alpha_{k,i}} \\
&\times \left(\gamma_2^{(k,i)} e^{\beta(\beta\sigma^2/2 - \overline{s}_{k,\ell}(\underline{\theta}))} \left(Q\left(\frac{a - \overline{s}_{k,\ell}(\underline{\theta}) + \beta\sigma^2}{\sigma}\right) - Q\left(\frac{\overline{s}_{k,j}(\underline{\theta}) - \overline{s}_{k,\ell}(\underline{\theta}) + \beta\sigma^2}{\sigma}\right)\right)\right)
\end{aligned}$$

6.2.3 Gaussian Noise with Random Deletions

For this noise model, each pixel is either kept or deleted, as determined by Bernoulli random variables associated with each pixel. If the pixel is not deleted, it is corrupted by additive K-D Gaussian noise with specified covariance. The kth pixel is distributed according to

$$p\left(\underline{s}_k\right) = \begin{cases} \frac{1-\alpha_k}{(2\pi)^{K/2}|\mathbf{\Sigma}_k|^{1/2}} e^{-\frac{1}{2}\left(\underline{s}_k - \overline{\underline{s}}^*_{k,0}\right)^T \mathbf{\Sigma}_k^{-1}\left(\underline{s}_k - \overline{\underline{s}}^*_{k,0}\right)} & \underline{s}_k \in \mathcal{A} \\ \frac{\alpha_k}{|\mathcal{B}|} & \underline{s}_k \in \mathcal{B} \end{cases} \tag{6.39}$$

where α_k is the probability that the kth pixel is deleted and $\mathbf{\Sigma}_k$ is the covariance matrix of each pixel.

For this noise type, the UIM can be used since the required expectations can be obtained analytically using the method described in Section 6.1.3 using (6.18) and the Gaussian noise results derived in Section 6.1.1 in (6.10).

The use of the random deletion model in this case is particularly attractive from an implementation standpoint because, as will be seen, it allows for the distribution of the image likelihood ratio to be determined analytically, unlike for the coherent LADAR model. As a result, calculations of error probability are much simpler and faster. In addition, the simplicity of the model lends itself to analytic approximations of performance as a function of system parameters, as will be seen.

6.2.4 Finite beamwidth effects

This section discusses performance prediction using the approximate model for finite beamwidth effects introduced in Section 3.2.4. The non-anomalous distribution is given by

$$p_{na}(\underline{s}_k|H_i) = \sum_{r=1}^{R_m^{(k,i)}} \frac{w_r}{\sqrt{(2\pi)^K |\mathbf{\Sigma}_r|}} e^{-\frac{1}{2}(\underline{s}_k - \underline{\mu}_r)^T \mathbf{\Sigma}_r^{-1} (\underline{s}_k - \underline{\mu}_r)} \tag{6.40}$$

where w_r is the associated probability for each portion, $\underline{\mu}_r$ is the mean, and $\underline{\Sigma}_r$ is the variance. This distribution can be used with the appropriate anomaly pdf p_a (deletion, uniform, or exponential) and probability of anomaly, giving

$$p(\underline{s}_k) = \alpha_k p_a(\underline{s}_k) + (1 - \alpha_k) p_{na}(\underline{s}_k) \tag{6.41}$$

The anomalous distribution p_a can be either that of coherent LADAR (uniform), the random deletion model, or of Geiger mode LADAR.

For the method of unbiased integration, the expectation

$$E\left[p\left(\underline{s}_k|H_i, \underline{\theta}\right)|H_0, \underline{\theta}_0^*\right] \tag{6.42}$$

is required. Since this model is a mixture of Gaussian pdfs and an anomalous pdf, the mixture pdf model results can be used along with the Gaussian noise results and the Gaussian noise plus anomalies results. From the mixture pdf results (6.16),

$$\begin{aligned}
&E\left[p\left(\underline{s}_k|H_i, \underline{\theta}\right)|H_0, \underline{\theta}_0^*\right] = \\
&\quad (1-\alpha_{k,i})(1-\alpha_{k,0}) \sum_{r_0=1}^{R_0} \sum_{r_i=1}^{R_1} w_{0,r_0} w_{i,r_i} E\left[p_{r_i}(\underline{s}_k|H_i, \underline{\theta})|p_{r_0}(\underline{s}_k|H_0, \underline{\theta}_0^*)\right] \\
&\quad + (1-\alpha_{k,i})\alpha_{k,0} \sum_{r_i=1}^{R_i} w_{1,r_1} E\left[p_{r_i}(\underline{s}_k|H_i, \underline{\theta})|p_a(\underline{s}_k|H_0, \underline{\theta}_0^*)\right] \\
&\quad + \alpha_{k,i}(1-\alpha_{k,0}) \sum_{r_0=1}^{R_0} w_{0,r_0} E\left[p_a(\underline{s}_k|H_i, \underline{\theta})|p_{r_0}(\underline{s}_k|H_0, \underline{\theta}_0^*)\right] \\
&\quad + \alpha_{k,i}\alpha_{k,0} E\left[p_a(\underline{s}_k|H_i, \underline{\theta})|p_a(\underline{s}_k|H_0, \underline{\theta}_0^*)\right]
\end{aligned}$$

where

$$p_{r_i}(\underline{s}_k|H_i,\underline{\theta}) = \frac{e^{-\frac{1}{2}(\underline{s}_k-\underline{\mu}_{r_i})^T\mathbf{\Sigma}_{r_i}^{-1}(\underline{s}_k-\underline{\mu}_{r_i})}}{\sqrt{(2\pi)^K|\mathbf{\Sigma}_{r_i}|}} \tag{6.43}$$

The expectations in the first sum in the expectation can be found using the Gaussian results in (6.10) and the expectations in the remaining terms are given in the derivations in (6.30), (6.18), or (6.34) depending on the anomaly model.

Chapter 7

Analytic Approximations for the Dependence of Performance on Major Parameters

7.1 Asymptotic Resolution Dependence for Independent Pixel Noise

In [2], Yen and Shapiro derive an approximation for the asymptotic behavior of the probability of error as a function of increasing FLIR sensor resolution, using the performance prediction presented in [1] and a Gaussian noise model. We use a similar method to obtain a somewhat more general result that is applicable to LADAR and can incorporate the more accurate prediction methods presented above.

We consider the case that the pixel density increases to infinity. First, let there be a reference pixel configuration consisting of $N = N_{ref}$ pixels. Now, let the number of pixels N vary. Assume that as N changes, the relative distribution of the densities of the pixels over the sensor focal plane or planes remains constant, and the region imaged does not change. In other words, the pixel spacings for the new configuration are chosen such that for each of the original N_{ref} pixels, exactly N/N_{ref}

of the new pixels in its immediate neighborhood can be uniquely assigned to them, to within one pixel (due to required quantization). For example, consider the case where the N_{ref} pixels are evenly spaced over a focal plane with spacing Δ_{ref}. Then, as N varied, the pixel spacing would remain even, with spacing $\Delta \approx \Delta_{ref}\sqrt{N_{ref}/N}$.

It is desired to approximate the behavior of the likelihood ratio as N varies while all noise and true nuisance parameters remain constant. Assume that the noise from pixel to pixel is independent regardless of N. Model the noiseless scene as viewed from the sensor with a continuum of pixels as being piecewise continuous. Assume also that the noise distribution parameters are only spatially dependent and vary piecewise continuously. The discontinuities, if they exist, would correspond to sharp boundaries, such as that between the target and the background. As $N \to \infty$, the differences between the values of pixels separated by a constant number of pixels approaches zero. The same is true for the noise parameters. Assuming that the likelihood function $p(\underline{s}_k|H_i,\underline{\theta})$ for a particular pixel is continuous with respect to the noise parameters and $\underline{s}_k$, we thus have that

$$\lim_{N\to\infty} p(\underline{s}^{(k_1)}|H_i,\underline{\theta}) = p(\underline{s}^{(k_2)}|H_i,\underline{\theta}) \tag{7.1}$$

with k_1 and k_2 defined such that k_1 remains in the same location and the location of k_2 is separated from that of k_1 by a constant number of pixel widths, all of which must be approaching zero by our assumptions. Further details as to this aspect of the derivation can be found in [2].

Based on the above discussion, we propose the approximation

$$\begin{aligned} p(\mathcal{I}(N)|H_i,\underline{\theta},N) &= \prod_{k=1}^{N} p(\underline{s}_k(N)|H_i,\underline{\theta},N) \approx \prod_{k=1}^{N_{ref}} \left(p(\underline{s}_k(N_{ref})|H_i,\underline{\theta},N_{ref})\right)^{N/N_{ref}} \\ &= p(I(N_{ref})|H_i,\underline{\theta},N_{ref})^{N/N_{ref}} \end{aligned} \tag{7.2}$$

thus ignoring quantization effects which may occur for N not an integer multiple of N_{ref}. This approximation is a generalization of the approximation in [2], in which the authors show that the approximation is asymptotically correct as $N_{ref}, N \to \infty$.

We now turn to apply this large N approximation to the various prediction methods in this thesis. All methods require the determination of $\underline{\theta}_1^*(N)$, which is the value of $\underline{\theta}_1^*$ determined at a

sensor resolution of N. The authors of [2] show that

$$\lim_{N\to\infty} \underline{\theta}_1^*(N) = \underline{\theta}_1^*(\infty) \tag{7.3}$$

so we make the asymptotic approximation that $\underline{\theta}_1^*(N) = \underline{\theta}_1^*(N_{ref})$. For the method of Grenander, we also make the assumption that the nuisance parameter priors are uniform, that is the $f_i(\underline{\theta})$ are constant. Using this, the threshold adjustment for the Laplace approximation method (4.12) is approximated as

$$\begin{aligned}
\nu_{laplace}(N) - \log\nu =& \frac{1}{2}\log\frac{\det\left(\ddot{E}_1(\overline{\mathcal{I}}_i^*, \underline{\theta}_1^*, \sigma, N)\right)}{\det\left(\ddot{E}_0(\overline{\mathcal{I}}_i^*, \underline{\theta}_0^*, \sigma, N)\right)} \\
\approx& \frac{1}{2}\log\frac{\frac{N}{N_{ref}}\det\left(\ddot{E}_1(\overline{\mathcal{I}}_0^*, \underline{\theta}_1^*, \sigma, N_{ref})\right)}{\frac{N}{N_{ref}}\det\left(\ddot{E}_0(\overline{\mathcal{I}}_i^*, \underline{\theta}_0^*, \sigma, N_{ref})\right)} \\
=& \nu_{laplace}(N_{ref}) - \log\nu
\end{aligned} \tag{7.4}$$

since (4.11)

$$E_i(\mathcal{I}, \underline{\theta}, \sigma, N) = -2\sigma^2\left(\log p(\mathcal{I}|H_i, \underline{\theta}, N) + \log f_i(\underline{\theta})\right) + c$$

and where dependence on N is shown explicitly.

For the Noiseless Integration Method and the Unbiased Integration Method, using (7.2), the approximate threshold adjustments are given by (4.15) (for arbitrary nuisance priors)

$$\begin{aligned}
\nu'(N)/\nu =& \frac{p(\overline{\mathcal{I}}_0^*(N)|H_1, \underline{\theta}_1^*, N)\int_S p(\overline{\mathcal{I}}_0^*(N)|H_0, \underline{\theta}, N) f_0(\underline{\theta}) d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*(N)|H_0, \underline{\theta}_0^*, N)\int_S p(\overline{\mathcal{I}}_0^*(N)|H_1, \underline{\theta}, N) f_1(\underline{\theta}) d\underline{\theta}} \\
\approx& \frac{p(\overline{\mathcal{I}}_i^*(N_{ref})|H_1, \underline{\theta}_1^*, N_{ref})^{N/N_{ref}}\int_S p(\overline{\mathcal{I}}_0^*(N_{ref})|H_0, \underline{\theta}, N_{ref})^{N/N_{ref}} f_0(\underline{\theta}) d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*(N_{ref})|H_0, \underline{\theta}_0^*, N_{ref})^{N/N_{ref}}\int_S p(\overline{\mathcal{I}}_0^*(N_{ref})|H_1, \underline{\theta}, N_{ref})^{N/N_{ref}} f_1(\underline{\theta}) d\underline{\theta}}
\end{aligned} \tag{7.5}$$

and (4.38)

$$\nu'(N)/\nu = \frac{E\left[p(\mathcal{I}(N)|H_1, \underline{\theta}_1^*, N)\right] \int_S E\left[p(\mathcal{I}(N)|H_0, \underline{\theta}, N)\right] f_0(\underline{\theta}) d\underline{\theta}}{E\left[p(\mathcal{I}(N)|H_0, \underline{\theta}_0^*, N)\right] \int_S E\left[p(\mathcal{I}(N)|H_1, \underline{\theta}, N)\right] f_1(\underline{\theta}) d\underline{\theta}} \tag{7.6}$$

$$\approx \frac{E\left[p(\mathcal{I}(N_{ref})|H_1, \underline{\theta}_1^*, N_{ref})\right]^{N/N_{ref}} \int_S E\left[p(\mathcal{I}(N_{ref})|H_0, \underline{\theta}, N_{ref})\right]^{N/N_{ref}} f_0(\underline{\theta}) d\underline{\theta}}{E\left[p(\mathcal{I}(N_{ref})|H_0, \underline{\theta}_0^*, N)\right]^{N/N_{ref}} \int_S E\left[p(\mathcal{I}(N_{ref})|H_1, \underline{\theta}, N_{ref})\right]^{N/N_{ref}} f_1(\underline{\theta}) d\underline{\theta}}$$

respectively. For these equations, it can be seen that to compute $\nu'(N)$ having already computed $\nu'(N_{ref})$, it is only necessary to take a power of the previously evaluated likelihood portions of the integrand and recompute the integral.

It remains to compute the pdf of the log-likelihood ratio $L' = \log(p(\mathcal{I}(N)|H_1, \underline{\theta}_1^*, N)) - \log(p(\mathcal{I}(N)|H_0, \underline{\theta}_0^*, N))$ given the true distribution (determined by H_0, $\underline{\theta}_0^*$) of the image (4.12). Since the case of large N is being considered and the pixel noises are independent, the pdf of L' asymptotically approaches a Gaussian as $N \to \infty$ by the Central Limit Theorem. This approximation is very good, since the convergence to a Gaussian pdf is almost always quite fast [9], and most LADAR images have large numbers of pixels. As a result, we need only obtain the mean and variance of L' to specify its pdf. From (7.2),

$$\begin{aligned}
E[L'(N)] =& E[\log(p(\mathcal{I}(N)|H_1, \underline{\theta}_1^*, N))|H_0, \underline{\theta}_0^*, N] \\
&- E[\log(p(\mathcal{I}(N)|H_0, \underline{\theta}_0^*, N))|H_0, \underline{\theta}_0^*, N] \\
=& E[\log(p(\mathcal{I}(N_{ref})|H_1, \underline{\theta}_1^*, N_{ref})^{N/N_{ref}})|H_0, \underline{\theta}_0^*, N_{ref}] \\
&- E[\log(p(\mathcal{I}(N_{ref})|H_0, \underline{\theta}_0^*, N_{ref})^{N/N_{ref}})|H_0, \underline{\theta}_0^*, N_{ref}] \\
=& \frac{N}{N_{ref}} \left(E[\log(p(\mathcal{I}(N_{ref})|H_1, \underline{\theta}_1^*, N_{ref}))|H_0, \underline{\theta}_0^*, N_{ref}]\right. \\
&\left. - E[\log(p(\mathcal{I}(N_{ref})|H_0, \underline{\theta}_0^*, N_{ref}))|H_0, \underline{\theta}_0^*, N_{ref}]\right) \\
=& \frac{N}{N_{ref}} E[L'(N_{ref})]
\end{aligned} \tag{7.7}$$

and

$$\begin{aligned}\mathrm{Var}[L'(N)] =&\mathrm{Var}[\log(p(\mathcal{I}(N)|H_1,\underline{\theta}_1^*,N))|H_0,\underline{\theta}_0^*,N]+\mathrm{Var}[\log(p(\mathcal{I}(N)|H_0,\underline{\theta}_0^*,N))|H_0,\underline{\theta}_0^*,N] \\ =&\frac{N}{N_{ref}}\left(\mathrm{Var}[\log(p(\mathcal{I}(N_{ref})|H_1,\underline{\theta}_1^*,N_{ref}))|H_0,\underline{\theta}_0^*,N_{ref}]\right. \\ &\left.+\mathrm{Var}[\log(p(\mathcal{I}(N_{ref})|H_0,\underline{\theta}_0^*,N_{ref}))|H_0,\underline{\theta}_0^*,N_{ref}]\right) \\ =&\frac{N}{N_{ref}}\mathrm{Var}[L'(N_{ref})]\end{aligned} \tag{7.8}$$

by the independence of the new pixels.

As a result, the approximate probability of error is given by

$$P(D_1|H_0,N)\approx Q\left(\frac{\nu'(N)-\frac{N}{N_{ref}}E[L'(N_{ref})]}{\sqrt{\frac{N}{N_{ref}}\mathrm{Var}[L'(N_{ref})]}}\right) \tag{7.9}$$

which, as we have seen, is asymptotically correct as $\sigma\to 0, N, N_{ref}\to\infty$.

To increase the simplicity of the prediction equation, in the case where the priors are uniform we propose using the threshold adjustment approximation for the method of Grenander (7.4) to obtain $\nu'(N)=\nu'(N_{ref})$ regardless of the method used to compute $\nu'(N_{ref})$. This is asymptotically valid since all the approximation methods are asymptotically equivalent as $\sigma\to 0$. In effect, this is equivalent to fitting Gaussian curves to the integrands used to compute $\nu'(N_{ref})$ in such a way that the value of the integrals, and thus $\nu'(N_{ref})$, are preserved. These Gaussian curves would then be used in the computation of the approximate $\nu'(N)$ using (7.5) or (7.6) which of course reduces to the desired $\nu'(N)=\nu'(N_{ref})$. This gives

$$P(D_1|H_0,N)\approx Q\left(\frac{\nu'(N_{ref})-\frac{N}{N_{ref}}E[L'(N_{ref})]}{\sqrt{\frac{N}{N_{ref}}\mathrm{Var}[L'(N_{ref})]}}\right) \tag{7.10}$$

If it is desired to obtain the error in the probability prediction inherent in the smoothness assumption for the image, it can be found by computing $E[L'(N)]$ and $\mathrm{Var}[L'(N)]$ and substituting in for the probability of error. This avoids having to recompute $\nu'(N)$, thus limiting the computational

expense. This error computation may be practicable in the case where the prediction is computed for a range of resolution values, since it would be computed at only a few points to verify the validity of the smoothness assumption for the curve.

7.2 Analytic Approximations for Gaussian Noise with Anomalies

In this section, we derive results for the Gaussian noise with anomalies model only. The simplicity of the Gaussian noise plus random deletions model lends itself to analytic analysis. Hence, as it is asymptotically equivalent to the coherent LADAR model, and a reasonable approximation of most anomaly models, we use the Gaussian noise with random deletions model for all derivations in this section. In this section, we have the goal of obtaining analytic or near analytic results for the variation of performance with various relevant parameters.

7.2.1 Probability of Anomaly Dependence

In this subsection, we apply the scaling properties of the probability of anomaly for the random deletion model derived in Section 6.1.3 to other anomaly models, such as the coherent LADAR model. We propose computing the mean, variance, and threshold of the pseudo log likelihood ratio test at a reference point α_{ref} for the accurate anomaly model using one of the accurate prediction methods, and then using this point to create an approximation of the performance curve that would result from varying the probability of anomaly α in the neighborhood of that point. This would be useful in determining what effects changes in the anomaly rate will have on performance, in order to aid rapid system design. We assume that the $\alpha_{i_0} = \alpha_{i_1}$. Using reference pseudo loglikelihood moments E_{ref}, σ_{ref}, and reference threshold ν'_{ref} computed for reference anomaly rate α_{ref}, we have the approximation, using (6.28), that the probability of error as a function of α is

$$P(D_1|H_0) \approx Q\left(\frac{\nu'_{ref} - \frac{1-\alpha}{1-\alpha_{ref}}E_{ref}}{\sqrt{\frac{1-\alpha}{1-\alpha_{ref}}\sigma^2_{ref}}}\right) \tag{7.11}$$

This provides a simple approximation to the variation of classification performance with anomaly rate.

Since for many cases the range swath is long enough and anomaly rate low enough that anomalies can be modeled as random deletions, this prediction is widely applicable to approximating performance variation as a function of anomaly rate for other anomaly based models, as well as for varying clutter levels. This type of analysis is important, since the anomaly rate is primarily determined by the receiver sensitivity, transmission power, and length of the range swath, as well as the clutter density, all of which are important system parameters.

7.2.2 Noise Covariance Dependence

First, we consider dependence on the Gaussian noise covariance, assuming that the covariances do not change between models. Let the parameter σ be such that all the terms of all covariance matrices are proportional to σ^2, that is $\mathbf{\Sigma}_k = \mathbf{S}_k\sigma^2$. We consider the variation of performance as σ varies. From Section 6.1.1, we have (6.13) that the Gaussian only pseudo log likelihood ratio has the following moments:

$$\begin{aligned} E[L'(\mathcal{I}, \underline{\theta}_0, \underline{\theta}_1)] &= -\sum_{j=1}^{N} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1})^T \frac{\mathbf{S}_k^{-1}}{\sigma^2} \overline{\underline{s}}^*_{k,0} \\ &\quad + \frac{1}{2}\sum_{j=1}^{N} \left(\overline{\underline{s}}_{k,0}{}^T \frac{\mathbf{S}_k^{-1}}{\sigma^2} \underline{s}_{k,0} - \overline{\underline{s}}_{k,1}{}^T \frac{\mathbf{S}_k^{-1}}{\sigma^2} \overline{\underline{s}}_{k,1} \right) \\ \mathrm{Var}[L'(I, \underline{\theta}_0, \underline{\theta}_1)] &= \sum_{j=1}^{N} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1})^T \frac{\mathbf{S}_k^{-1}}{\sigma^2} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1}) \end{aligned}$$

For simplicity of notation, let

$$\begin{aligned} \mu_L &= -\sum_{k=1}^{N} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1})^T \mathbf{S}_k^{-1} \overline{\underline{s}}^*_{k,0} \\ &\quad + \frac{1}{2}\sum_{k=1}^{N} \left(\overline{\underline{s}}_{k,0}{}^T \mathbf{S}_k^{-1} \overline{\underline{s}}_{k,0} - \underline{s}_{k,1}{}^T \mathbf{S}_k^{-1} \overline{\underline{s}}_{k,1} \right) \\ \sigma_L^2 &= \sum_{k=1}^{N} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1})^T \mathbf{S}_k^{-1} (\overline{\underline{s}}_{k,0} - \overline{\underline{s}}_{k,1}) \end{aligned} \tag{7.12}$$

Furthermore, since the non anomalous Gaussian loglikelihood is inversely proportional to the Gaussian covariance, which is in turn proportional to σ^2, its Hessian with respect to the nuisance parameters must also be proportional to σ^2. Since this is true for both target loglikelihoods, the threshold adjustment $\nu_{laplace}$ for the method of Grenander does not change with σ. Hence, we assume that the threshold adjustment is constant regardless of prediction method, in a similar way as done in Section 7.1. To improve accuracy, we use a reference point σ_{ref} for which the performance has been predicted using one of the threshold adjustment methods, giving a threshold ν'_{ref}.

Using the results from Section 6.1.3, we then have, assuming the anomaly rates P_a do not vary with k, the approximate probability of error to be (7.13)

$$P(D_1|H_0) \approx Q\left(\frac{\nu'_{ref} - (1-\alpha)\frac{\mu_L}{\sigma^2}}{\sqrt{(1-\alpha)\frac{\sigma_L^2}{\sigma^2}}}\right) \tag{7.13}$$

which explicitly gives an analytic expression for the variation of performance with both anomaly rate and Gaussian noise level, since $\nu'_{ref}, \mu, \sigma_L$ are all constants. This noise level variation result is clearly also applicable for nonuniform anomaly rates, but is not shown for clarity. As the only approximations required for this result are that the threshold does not vary with σ and that the Central Limit Theorem can be used, this result is asymptotically correct as the noise level goes to zero and the number of pixels increases to infinity, assuming relevant Central Limit requirements are met. The first result is due to the approximation's basis in the method of Grenander, which is asymptotically accurate. The latter result is due to the requirements for the asymptoticity from the Central Limit Theorem and its extensions.

7.2.3 Elevation Angle Dependence

Next, we derive approximate variation results for sensor elevation angle ϕ for range images only, on a flat ground plane. Define the elevation angle to be the angle between the sensor line of sight and the ground plane. It is necessary to obtain a simple expression for the variation of the statistics of the pseudo loglikelihood ratio. For the Gaussian noise with random deletions model, these only depend on the variation of the distance between the predicted pixel values under the different hypotheses.

For most ground targets viewed from the air, the visible surfaces of the targets tend to be parallel to the ground plane. In addition, the dominant pixels from a separability standpoint will be those with the maximum difference between the hypothesis predictions, which will usually occur where one hypothesis says that the pixel is on the top of the target, and the other that it is on the ground plane. As the top of the target is usually parallel to the ground, this suggests that we approximate the average variation of the difference in hypothesized values for the pixels as the two model surfaces are locally horizontal. As the elevation angle varies, then, this would mean that the difference in pixel hypothesis values would vary as $1/\sin\phi$, using the far field approximation. That is,

$$\frac{1}{N_{targ}}\sum_{j=1}^{N}(\overline{s}_{k,0}^{(\phi)}-\overline{s}_{k,1}^{(\phi)})^2 \approx \frac{\sin^2\phi_{ref}}{\sin^2\phi}\frac{1}{N_{targ,ref}}\sum_{k=1}^{N}\overline{s}_{k,0}^{(\phi_{ref})}-\overline{s}_{k,1}^{(\phi_{ref})})^2 \tag{7.14}$$

where $N_{targ,ref}$ and N_{targ} are the numbers of pixels for which the pixel difference is nonzero for ϕ_{ref} and ϕ respectively. As the variation in loglikelihood is approximated to be due entirely to a single multiplicative constant, the threshold adjustment $\nu_{laplace}$ found using Grenander's method does not change. Hence, we approximate the threshold as constant with respect to elevation angle.

It remains to determine the variation of the number of pixels on target, that is, the number for which the difference in hypothesized values is nonzero. We consider low elevation angles and high elevation angles separately. For low elevation angles, portions of both the side and top of the target are visible. Hence, it is reasonable to approximate the number of pixels as unchanging with elevation angle. This approximation is exact for a cylindrical target. As a result, this gives the variation of the moments of the pseudo LLR with zero probability of anomaly to be

$$\begin{aligned} E[L'|\phi] &\approx \frac{\sin^2\phi_{ref}}{\sin^2\phi}E[L'|\phi_{ref}] \\ \mathrm{Var}[L'|\phi] &\approx \frac{\sin^2\phi_{ref}}{\sin^2\phi}\mathrm{Var}[L'|\phi_{ref}] \end{aligned} \tag{7.15}$$

This gives an approximate probability of error for arbitrary probability of anomaly to be, using the

Central Limit Theorem

$$P(D_1|H_0) \approx Q\left(\frac{\nu'_{ref} - (1-\alpha)\frac{\mu_{ref}}{\sigma^2}\frac{\sin^2\phi_{ref}}{\sin^2\phi}}{\sqrt{\left((1-\alpha)\frac{\sigma^2_{ref}}{\sigma^2}\frac{\sin^2\phi_{ref}}{\sin^2\phi}\right)}}\right) \tag{7.16}$$

where

$$\mu_{ref} = -\sum_{j=1}^{N}(\overline{s}_{k,0}^{(\phi_{ref})} - \overline{s}_{k,1}^{(\phi_{ref})})\overline{s}_{k,0}^{(*,\phi_{ref})} + \frac{1}{2}\sum_{j=1}^{N}\left(\overline{s}_{k,0}^{(\phi_{ref})^2} - \overline{s}_{k,1}^{(\phi_{ref})^2}\right) \tag{7.17}$$

$$\sigma_r ef^2 = \sum_{k=1}^{N}(\overline{s}_{k,0}^{(\phi_{ref})} - \overline{s}_{k,1}^{(\phi_{ref})})^2$$

and ν'_{ref} is computed at ϕ_{ref}.

For high elevation angles, the majority of the image of the target is of the flat top surface. Hence, the variation in the number of pixels on target can be approximated by the variation in the apparent area of a horizontal flat plate, using the far field approximation. This approximation is asymptotically correct as the height of the target goes to zero. This indicates that the number of pixels on target should be proportional to $\sin\phi$. This clearly gives the approximate probability of error

$$P(D_1|H_0) \approx Q\left(\frac{\nu'_{ref} - \left((1-\alpha)\frac{\mu_{ref}}{\sigma^2}\frac{\sin\phi_{ref}}{\sin\phi}\right)}{\sqrt{(1-\alpha)\frac{\sigma^2_{ref}}{\sigma^2}\frac{\sin\phi_{ref}}{\sin\phi}}}\right) \tag{7.18}$$

Experiments will be run in a later section to verify the accuracy of these approximations.

Chapter 8

Bayesian Experimental Results for LADAR Noise

In this chapter, we present experimental results for the three prediction methods discussed for each of the noise models considered. Various binary sets of targets are used. The nuisance parameter for each of these experiments was target azimuth angle for purposes of illustration.

Each experiment involves the computation of Monte Carlo simulation results for the actual Bayesian classifier, and the evaluation of the Method of Noiseless Integration and Method of Unbiased Integration predictions. For the Gaussian range noise model, the method of Grenander is also used for comparison. Performance curves are generated for a relevant range of noise levels in each case, making sure that the curve includes both high and low performance regions. The accuracy of the predictions is then evaluated and compared for each noise type.

8.1 1-D Gaussian Noise

For this section and some following, we use the tank from Figure 4.1 as target H_0, and the sedan from Figure 4.2 as target H_1.

An experiment was performed to evaluate the relative performance of each of the methods described in this paper. A binary classification scenario consisting of two targets on a flat ground plane

being imaged by a LADAR sensor was considered. The LADAR image was modeled as being a 2-dimensional range image of the scene, corrupted by one-dimensional IID Gaussian pixel range noise of zero mean and variance σ^2. The two targets are a tank and a sedan, shown in Figures 4.1 and 4.2 respectively. The tank corresponds to the true hypothesis H_0 and the sedan to the alternative hypothesis H_1. Synthetically generated noiseless LADAR imagery was used in the evaluation of the prediction formulas derived above. This was accomplished using the GTRI LADAR Simulator software [3], which is able to efficiently generate synthetic noiseless LADAR imagery for any desired set of pose parameters. The only nuisance parameter considered was angular pose with a uniform prior from $-\pi/2$ to $\pi/2$ for numerical convenience. The predictions of the probability of error given H_0 is true for each prediction method as a function of σ/d_{min} are shown in Figure 8.1, where d_{min} denotes the minimum root sum squared error over the nuisance parameter space between the true noiseless image and the target for H_1. Normalization by d_{min} gives the minimum statistical distance resulting from using σ, thus providing a sort of dominant SNR. For this set of targets, $d_{min} = 0.18$. The prediction using the method of Grenander is shown in red, that using the NIM is shown in green, and that using the UIM is shown in black. The method of Grenander is shown only for this noise type because is was originally derived only for Gaussian noise. For evaluation of the accuracy of these methods, a plot of the empirical performance of the Bayesian classifier is also shown in blue. The empirical performance was found using 1 million Monte Carlo runs for each noise level, using synthetically generated imagery.

It can be seen from the plot that, as expected, it appears that all three methods are asymptotically accurate as $\sigma \to 0$, and are less accurate as the noise increases. As might be expected, the method of Grenander fails for high noise. In opposition to this, the two new methods do not diverge from the truth as the noise increases. The NIM performs worse of the two new methods, but significantly better than the method of Grenander. In particular, it remains within approximately 0.1 of the true probability of correct classification for all noise levels. The method with the best performance is the UIM. It is better than the NIM at every noise level, and remains within 0.01 of the truth for all but the highest noise levels.

An important aspect of Bayesian classification is that for equal priors the conditional error

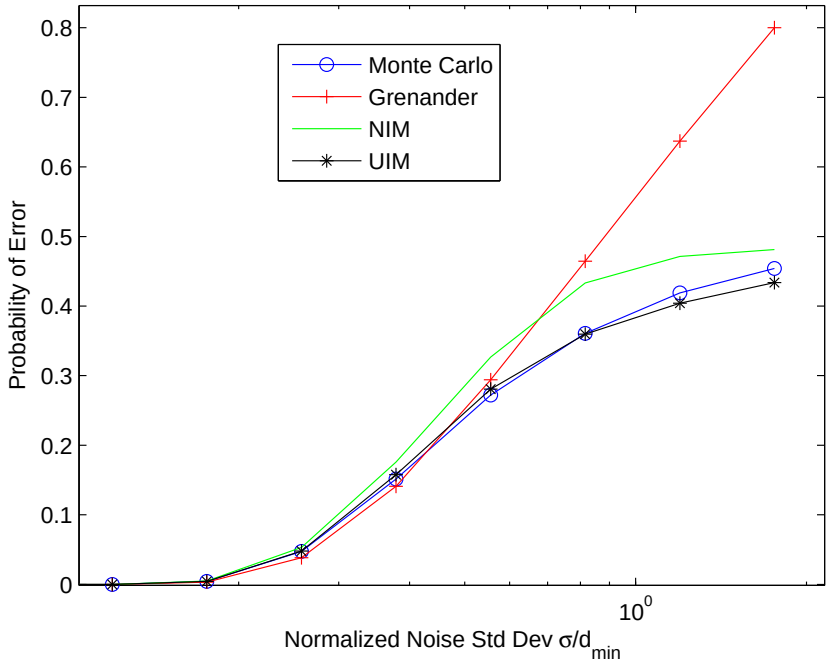

Figure 8.1: Probability of error given H_0 is true as a function of normalized noise standard deviation σ/d_{min}. Monte Carlo simulation results shown in blue, method of Grenander shown in red, NIM shown in green, and UIM shown in black.

probabilities saturate at 50% as the noise level increases without bound. As is seen in the figure, the prediction based on the method of Grenander does not exhibit this behavior, while the two methods introduced here do. This ensures that our methods will be accurate on both sides of the noise curve, whereas the method of Grenander is only accurate on the low-noise side. Note also that the NIM and the UIM differ considerably for high noise, but not as much for low noise. This indicates that the accurate choice of ν' is much more critical for prediction in high noise than in low [2].

8.2 1-D Gaussian with Anomalies- Coherent

For this experiment, we use the Tank and Civic as targets, and use the noise model of Section 6.2.1. For this scenario, the probability of anomaly was chosen to be 0.2. Both the Unbiased Integration and Noiseless Integration Methods are used to predict the performance. To calculate the required LLR pdf for the predictions, we use numeric integration. The Monte Carlo results were generated using 2×10^4 runs. The results are shown in Figure 8.2.

As can be seen, the UIM performs much better than does the NIM. The NIM performs reasonably well for lower noise levels, but then becomes highly pessimistic for high noise, with error peaking around .13. The UIM is accurate to within approximately .03 in absolute probability of error for all noise levels, performing slightly worse than for Gaussian noise alone but still quite good. Since the Laplace approximation method must always be worse than the NIM, this experiment confirms that the UIM is the most accurate prediction method for this problem.

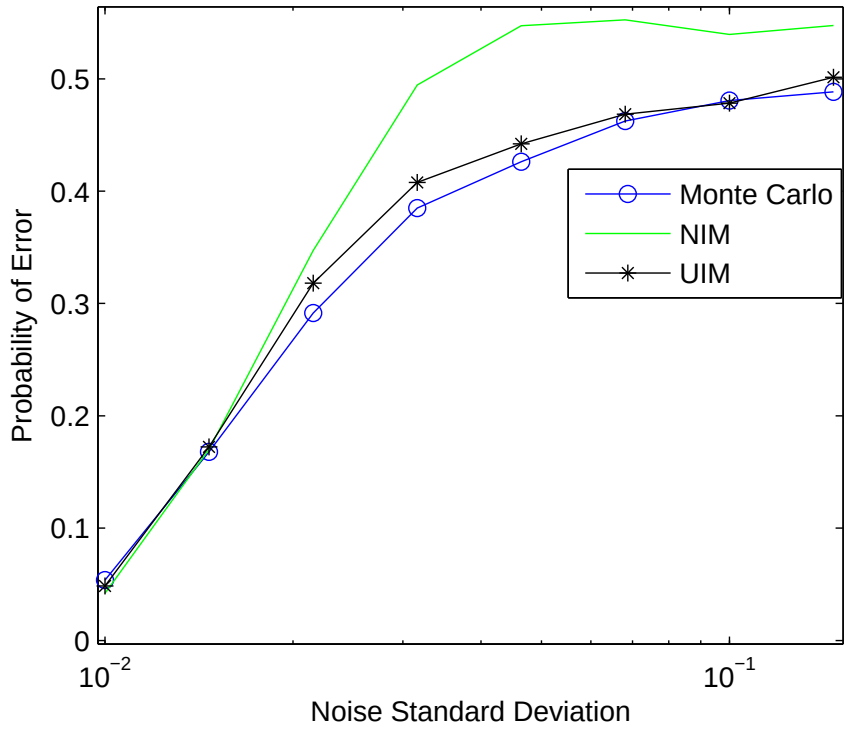

Figure 8.2: Performance prediction for coherent LADAR noise model. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

8.3 1-D Gaussian Noise with Geiger-mode Anomalies

In this section, we show experimental results for the Gaussian noise with Geiger mode LADAR false alarms model. The targets are the Civic in 9.1 and Avalon in 9.2, with the true target being the Civic. The nuisance parameter is azimuth angle, with a uniform prior over $-\pi$ to π. Parameters are chosen

such that the overall probability of detection is approximately 0.75, the probability of deletion 0.07, the probability of early anomaly approximately 0.16, and the probability of late anomaly 0.016. The early and late range swaths are approximately of equal length. The probability of a received photon at the correct range is 0.9, hence the result that the late anomaly probability is about one tenth that of the early anomaly probability. This causes a discontinuity in the pdf. For this experiment, we vary the Gaussian noise standard deviation only while maintaining the anomaly distribution, with the range swath long enough that the Gaussian noise distribution does not begin "overflowing." As a result, as the noise level increases, the relative height of the anomaly pdf relative to the Gaussian pdf increases, thus accentuating the effect of the pdf discontinuity, as will be seen.

The results showing the probability of error as a function of Gaussian noise standard deviation are shown in Figure 8.3. The Monte Carlo results using 1×10^4 simulations are shown in blue, the Noiseless Integration prediction in green, and the UIM prediction in black. The NIM is quite poor, however. This is because the integrand is evaluated at the noiseless value of the image, which is exactly where the pdf discontinuity lies. Hence as the nuisance parameter varies slightly, large numbers of pixels move across the discontinuity, thus causing large spikes in the integrand and destroying the prediction. It can be seen, on the other hand, that the UIM prediction performs quite well. This is due to the fact that it does not evaluate the likelihood function at a particular image value, but instead computes an expectation, which is of course not subject to the discontinuity effect which damages the prediction using the NIM. This indicates that the UIM is not only more accurate, but more robust.

8.4 1-D Gaussian Noise with Random Deletions

For this experiment, the model of Gaussian noise with random deletions was used, with the Tank and Civic as targets and with probability of deletion of .2. The plot of the probability of correct classification as a function of the Gaussian noise standard deviation is shown in Figure 8.4. The Monte Carlo truth plot was obtained using 2×10^4 Monte Carlo runs for each noise level. As can be seen, the Unbiased Integration Method prediction is again quite good, and outperforms that of the Noiseless Integration Method, especially for high noise levels. As before, the Noiseless

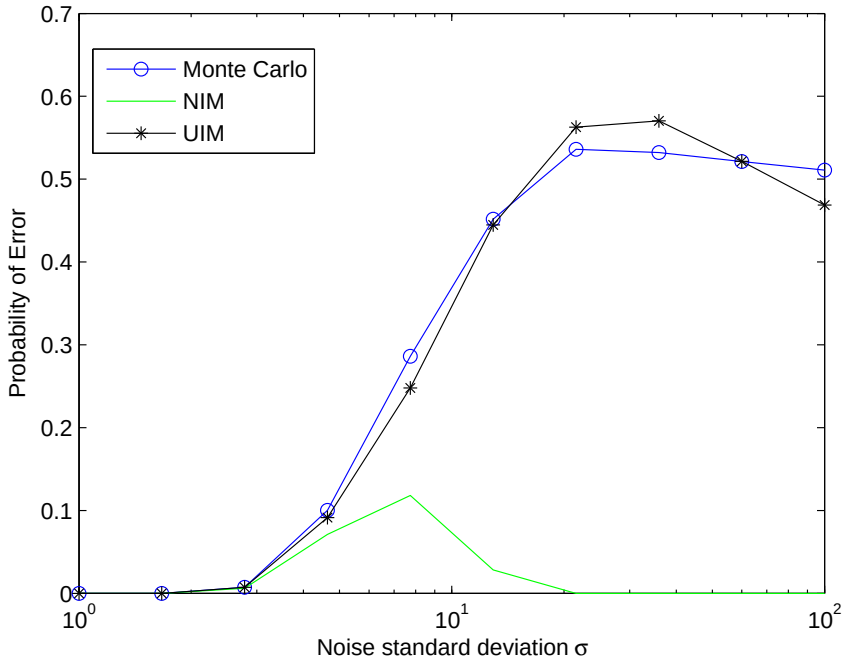

Figure 8.3: Performance prediction for Geiger LADAR noise model. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, Method of Noiseless Integration shown in green, and Method of Unbiased Integration shown in black. Note the poor performance of the Noiseless Integration prediction due to the discontinuity in the pdf.

Integration Method prediction works well for low noise, but then decays too rapidly to above 0.5, before flattening out. On the other hand, the prediction of the Unbiased Integration Method tracks the true Monte Carlo curve quite well, with a maximum error of around 0.02.

8.5 3-D Gaussian Noise with Random Deletions

For this experiment, 3-D point cloud images were used with the 3-D Gaussian noise plus random deletions model. The nuisance parameter is target angle with a uniform prior between $-\pi/2$ and $\pi/2$ as before. The 3-D noise covariance was chosen to be

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & 4\sigma^2 \end{bmatrix} \tag{8.1}$$

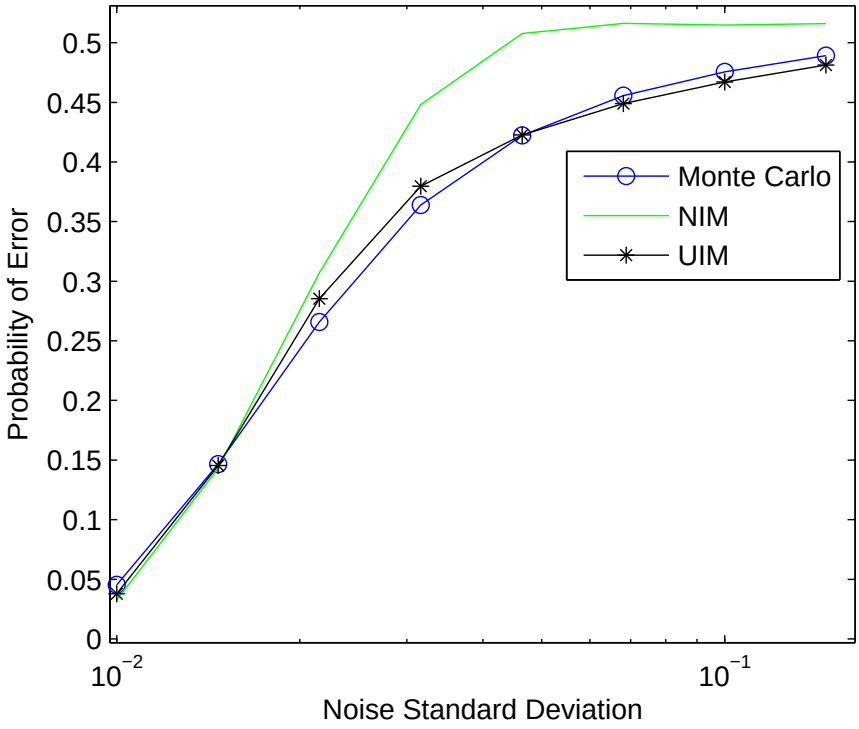

Figure 8.4: Performance prediction for Gaussian noise with random deletions. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, Method of Noiseless Integration shown in green, and Method of Unbiased Integration shown in black.

where σ in this case is chosen so that one-dimensional plots as a function of noise level can be presented. The probability of detection P_a was chosen to be 0.8 as before. The results are shown in Figure 8.5. The Monte Carlo truth plot was generated using 10^4 simulations per noise level. As expected, the prediction using the Unbiased Integration Method is the best, with a maximum error of 0.012. The prediction using the Noiseless Integration Method is worse, but not by as large a margin as for range images, with a maximum error of 0.04. The good performance of the Noiseless Integration Method is likely because the three dimensional noise greatly reduces the sharp variations of the likelihood functions as functions of pose, thus making the problem easier to work with and predict.

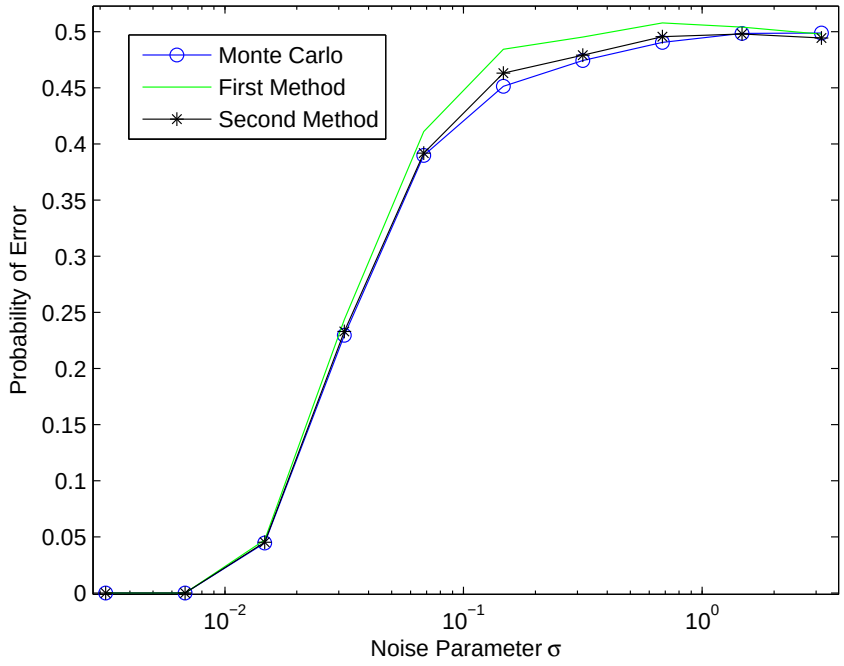

Figure 8.5: Performance prediction for 3-D Gaussian point cloud noise with random deletions. Probability of error given H_0 is true as a function of Gaussian parameter σ. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

8.6 Coherent LADAR with finite beamwidth effects

In this section, we evaluate the performance prediction accuracy for the finite beamwidth effects noise model. The targets are the Honda Civic 9.1 and Toyota Avalon 9.2. The nuisance parameter is target orientation, with a uniform prior on $-\pi/2$ to $\pi/2$. We use the coherent LADAR uniform anomaly model, with an anomaly rate of 0.2. The image size is 50×50 pixels. The beamwidth of the receivers is chosen to have a half power width of about one pixel width. To implement the finite beamwidth effect, we use a raised Gaussian beam profile and sample the 25 surrounding pixels in the 100×100 pixel model images, corresponding to a width of 2 pixels in the actual image, thus allowing for the beam profile to decay from its half power width. The use of the higher resolution image allows for greater accuracy in the discretization of the range pdf. For each pixel, the probability that each of the 25 range values is chosen is computed using the associated power density for that subpixel. Each of these range values is then used as the mean of a weighted Gaussian pdf to create a Gaussian mixture model, all with a standard deviation σ as described in the model.

The results are shown in Figure 8.6, with the Monte Carlo results shown in blue, the Noiseless Integration prediction in green, and the Unbiased Integration prediction in black. The prediction accuracies are typical, with the Unbiased Integration performing quite well, and much better than the Noiseless Integration method.

For comparison, we also show the prediction results for zero beamwidth noise in Figure 8.7. It can be seen that the performance is signficantly better than for the finite beamwidth case for low noise. This is as expected, since it involves less uncertainty in range. As the noise level increases, the performance difference decreases due to the fact that the difference in the possible ranges becomes less significant due to the very large noise power.

Finally, we show the results for the case where the true noise is distributed according to the finite beamwidth effects, but the classifier incorrectly uses the zero beamwidth noise model with the same value of σ. This is an example of the common case that, for speed, the classifier uses a simple noise model, but the actual noise is known to be much more complex. The results are obtained using the method of Section 4.7.3 and are shown in Figure 8.8. The Unbiased Integration prediction is quite good. The Noiseless Integration method performs significantly worse for this scenario. This is also as expected given that the Unbiased Integration method lends itself much more easily to the task of prediction performance under incorrect noise models.

The performance for the incorrect noise model is significantly worse than that for when the correct finite beamwidth model is used, especially for low noise. This is as expected, since for high noise, the different range values become statistically closer, thus reducing the effect of the finite beamwidth model. This shows that using too simple of a noise model for the classifier can cause a major degradation in performance, as most systems attempt to operate in the very low error rate region. It also indicates that accurate performance prediction for the low noise regime depends much more on the accuracy of the imaging model than on the choice of prediction method.

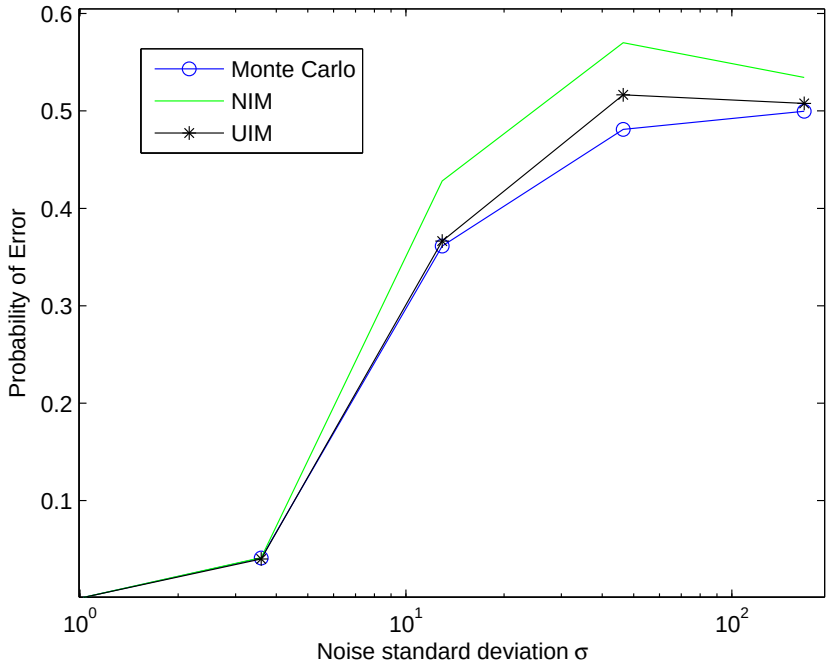

Figure 8.6: Performance prediction for finite beamwidth noise with uniform anomalies. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, Method of Noiseless Integration shown in green, and Method of Unbiased Integration shown in black.

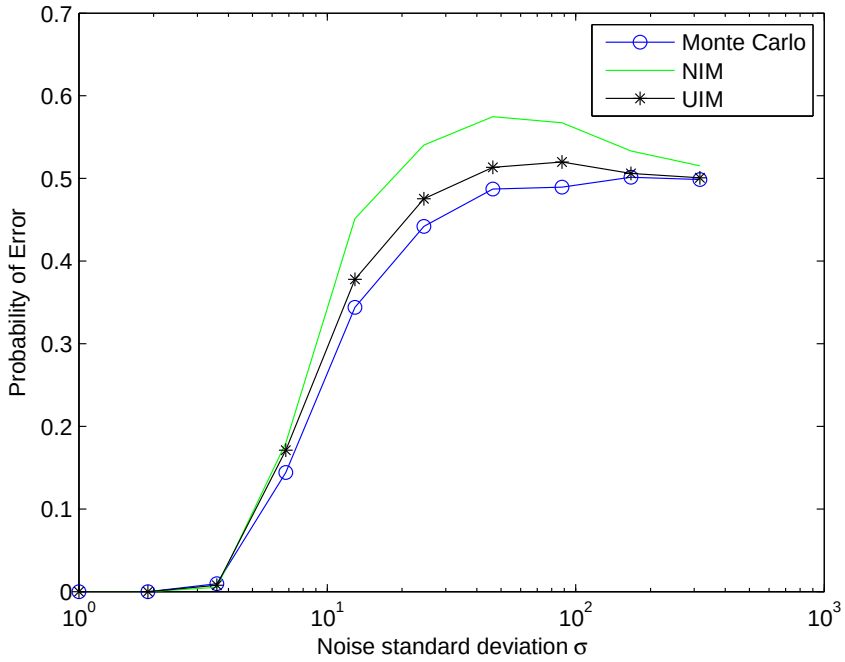

Figure 8.7: Performance prediction for zero beamwidth noise with uniform anomalies, for comparison to finite beamwidth results. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, Method of Noiseless Integration shown in green, and Method of Unbiased Integration shown in black.

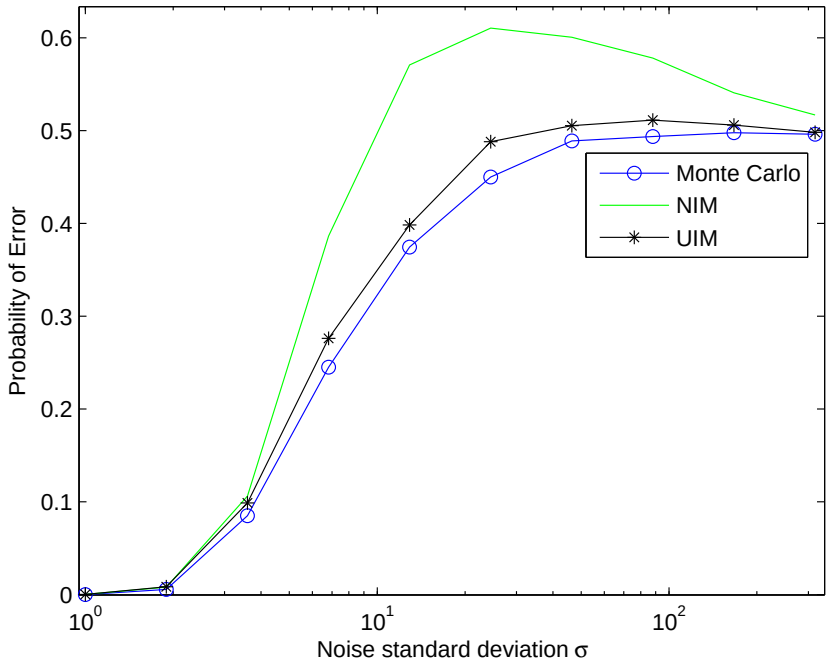

Figure 8.8: Performance prediction for finite beamwidth noise with uniform anomalies. Classifier uses incorrect zero beamwidth noise model. Probability of error given H_0 is true as a function of Gaussian noise standard deviation σ. Monte Carlo simulation results shown in blue, Method of Noiseless Integration shown in green, and Method of Unbiased Integration shown in black.

Chapter 9

Experimental Studies

In this section, we turn to the use of the prediction methods developed in this thesis for predicting performance as functions of various parameters. In addition, we verify the accuracy of the various extensions to the main methods described above, including the multimodal extension, the multiple target extension, prediction using multiple nuisance parameters, and the various derived asymptotic expressions for the variation of performance with different parameters.

In this chapter, we primarily use the targets shown in Figures 9.1 and 9.2, which are a Honda Civic and Toyota Avalon respectively, shown in noiseless range images of 200×200 pixels for clarity. This additional set of targets is introduced to demonstrate that these methods work on a variety of targets. These targets are also somewhat more realistic than the previous sedan/tank pair, in that they have a generally more similar appearance. Unless otherwise specified, the true target is always the Honda Civic, in a pose pointing away from the sensor.

9.1 Performance for Multiple Target Classification

In this section, we verify the accuracy of our performance prediction methods for the multiple target case. We consider a four target ATR system, with the true target being the Avalon in Figure 9.2, and the other three targets being the Civic shown in Figure 9.1, a Mazda Sentra shown in Figure 9.3, and a Jeep in Figure 9.4. Images of size 56×56 pixels are used. The nuisance parameter is

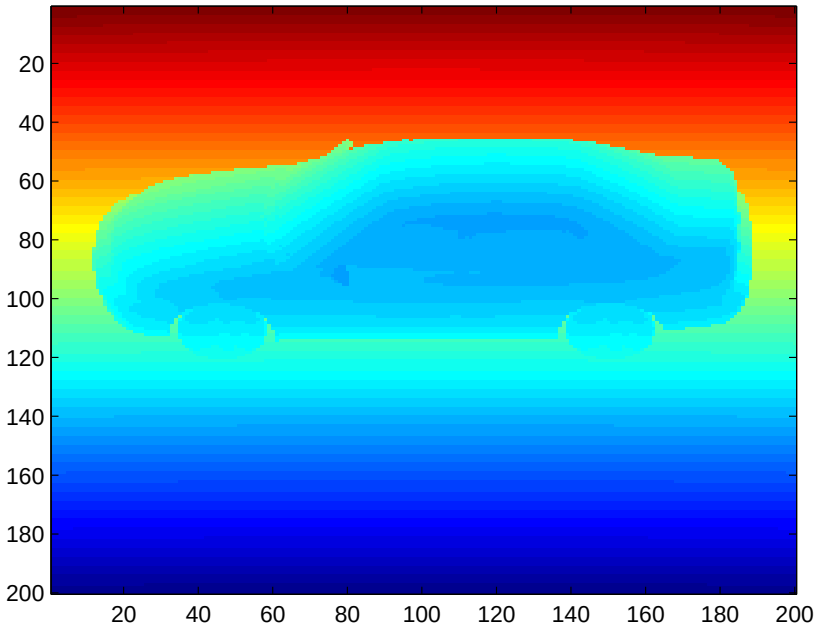

Figure 9.1: Honda Civic Model, shown in a noiseless LADAR range image.

target azimuth angle, with prior $-\pi$ to π radians. The noise model is Gaussian range noise. The results are shown in Figures 9.5-9.8. Monte Carlo results were generated using 3×10^4 simulations and are shown in blue. The Noiseless Integration prediction is shown in green, and the Unbiased Integration prediction in black. The plot of the probability that the true Avalon target is selected is shown in Figure 9.5. Note the initial large probability of correct classification, which decays to near zero as the noise level increases. This is because the Avalon is sandwiched between two targets, the Civic and Sentra, one of which is slightly larger, and one slightly smaller. This results in the thresholds determining the decision rules to be drawn so as to hem in the true target, and thus reduce its probability of correct classification as the noise becomes high.

The probability that the Civic is chosen is shown in Figure 9.6. Note that the probability increases rapidly from zero, before decreasing again as the noise becomes very large. This is expected, since the Civic is similar to the Avalon, and in terms of size between the Avalon and Jeep. The probability that the Sentra is chosen is shown in Figure 9.7, and that for the Jeep in 9.8. As can be seen, the probability of the Jeep being chosen is initially very low, and remains so for low noise. This is as expected, since the Jeep is the target most unlike the truth. For this region, the

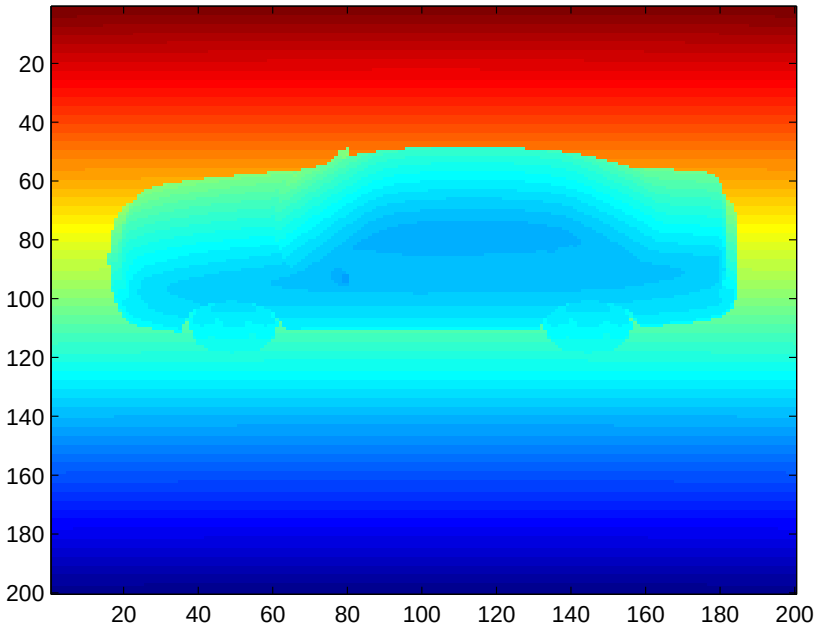

Figure 9.2: Toyota Avalon Model, shown in a noiseless LADAR range image.

probability of error is dominated by the two targets closest to the truth, the Sentra and Civic. As the noise becomes very high, however, the probability the Jeep is chosen begins to increase. Again, this is as expected since the increased noise hides the dissimilarities of the targets. This increase comes at the expense of the other three targets, but especially of the Civic. This is as expected, since both the Civic and Jeep are larger than the true target, whereas the Sentra is somewhat smaller. Hence the Civic is allotted less of the space and loses its probability as the noise decreases the distance between targets.

The prediction methods do not perform quite as well as with the binary case, but this is to be expected given the larger number of degrees of freedom for the predicted quantities. Overall, though, the predictions are still reasonably good, especially that of the Method of Unbiased Integration. As always, the accuracy is greatest for low noise, where the prediction matters most. If more accuracy is needed, the multimodal method of Section 4.7.1 can be used to achieve arbitrary accuracy by increasing the number of approximating random samples.

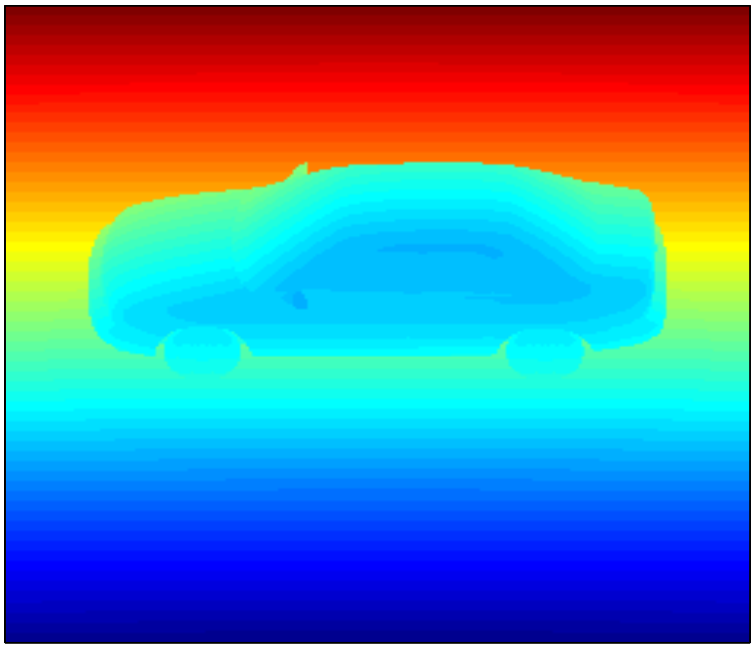

Figure 9.3: Mazda Sentra Model, shown in a noiseless LADAR range image.

9.2 Performance for Multiple Nuisance Parameters

In this section, we consider performance prediction in the presence of multiple nuisance parameters. The Civic and Avalon targets are used with image size 56×56 pixels, with the Civic being the true target. The three common nuisance pose parameters [3] are used, which are (x, y) position and azimuth angle. The prior on the azimuth angle is uniform from $-\pi/2$ to $\pi/2$. The other priors are uniform as well. Since 3-D integration is required, we use 11 by 11 by 19 grids of points surrounding the main peak. The range of these grids are chosen using the optimization method described in Chapter 5. Numeric integration is performed using Gaussian interpolation as described in Chapter 5. The Gaussian noise with random deletions model is used, with probability of deletion 0.2.

The results are shown in Figure 9.9, with the Monte Carlo results shown in blue, the Noiseless Integration prediction in green, and the Unbiased Integration prediction in black. As expected, the UIM prediction outperforms the NIM prediction. It can also be seen that the predictions are slightly

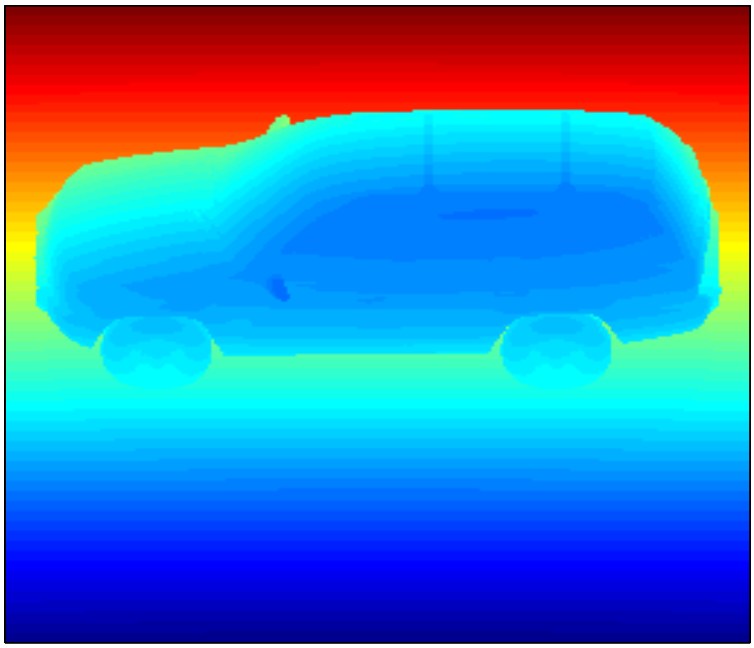

Figure 9.4: Jeep Model, shown in a noiseless LADAR range image.

optimistic. This is to be expected due to the greater inaccuracy with the 3-D integration.

9.3 Target Symmetry

In this section, the performance of classifying the Civic and Avalon is considered, with the nuisance parameter being orientation angle. The true target is assumed to be the Civic in Figure 9.1. We use images of size 56×56. In this scenario, the targets have a similar appearance after a 180 degree rotation, thus making the necessary likelihood integrals used for recognition (4.1) and the calculation of ν' bimodal.

We perform this experiment with the coherent LADAR model, with a probability of anomaly of 0.2. The results are shown in Figure 9.10. The Monte Carlo simulation results using 10000 simulations for the full width prior from $-\pi$ to π are shown in blue, the prediction using the NIM is shown in green, and the UIM prediction is shown in black. The results using the half prior of

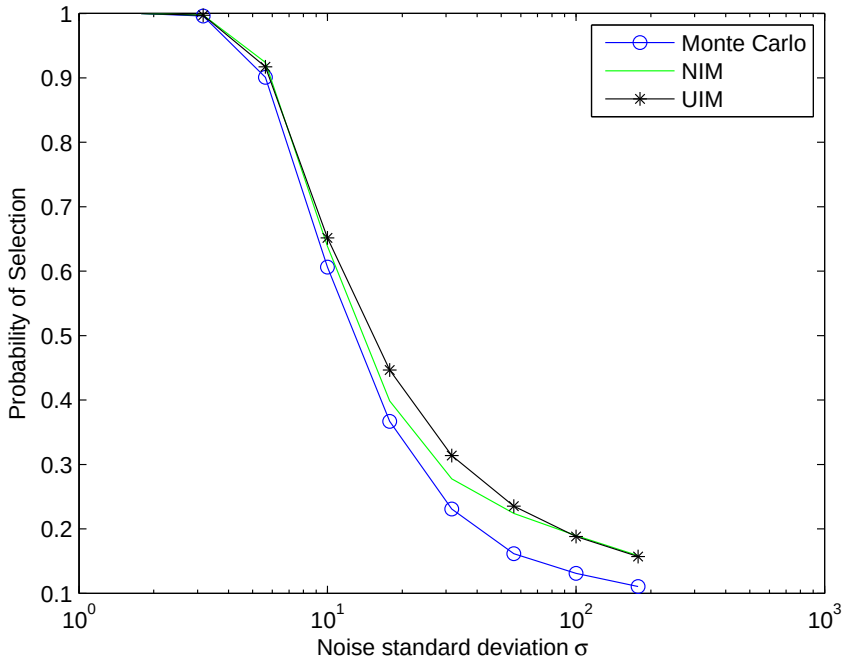

Figure 9.5: Performance prediction for 1-D Gaussian range noise. Probability that the correct target (Avalon) is chosen as a function of noise standard deviation. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black. Note the decay from perfect classification to very poor as noise increases.

$-\pi/2$ to $\pi/2$ are also shown, with Monte Carlo simulation results shown in red, the NIM prediction shown in magenta, and the UIM prediction shown in cyan. Note the degradation in performance in moving to the more realistic prior from the half width prior. This is because the presence of symmetries makes the recognition problem more difficult. Both prediction methods work well, with the Unbiased Integration performing much better than the Noiseless Integration method, and the prediction being slightly better for the half width prior due to the lack of symmetries. From the difference in performance between the true priors it appears that if the method of Grenander were as accurate as possible, the prediction should result in predictions that are overly optimistic by at least 0.1. This is because the Laplace approximation assumes that the integrand has a single peak, and thus at best can only integrate over the region near the primary peak.

Since the functions to be integrated are bimodal, we expect the multi-modal extension of the prediction methods in Section 4.7.1 to improve the prediction. We test this hypothesis in the next section.

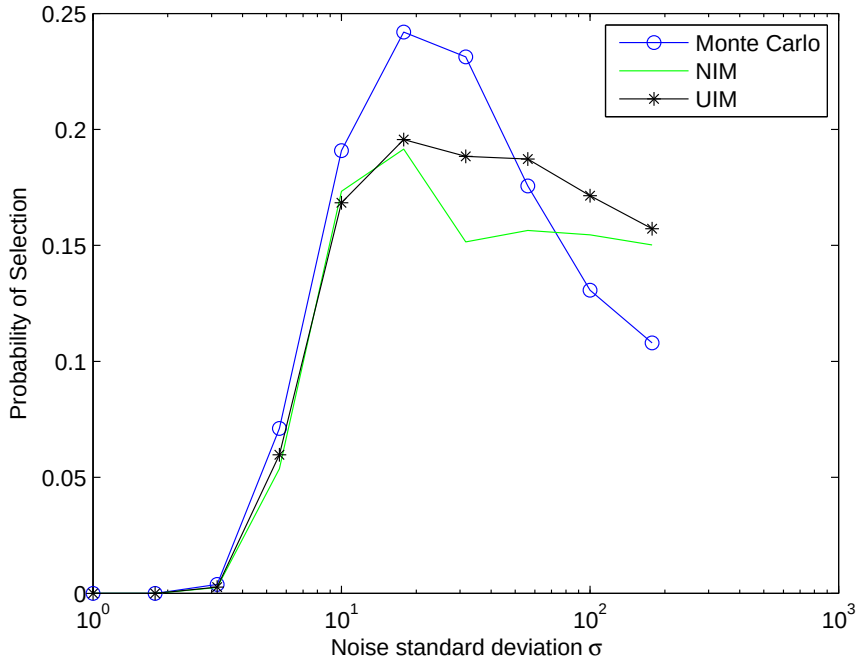

Figure 9.6: Performance prediction for 1-D Gaussian range noise. Probability that the Civic is chosen as a function of noise standard deviation. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black. Note the rapid increase in probability followed by a decay.

9.4 Multi-Modal Extension

In this section, the use of the multi-modal extension is demonstrated. The noise model of Gaussian noise plus random deletions is used, with probability of deletion 0.2. The ATR problem is the same as the full ($-\pi$ to π) prior scenario in the previous section. In that section, it was found that the prediction using one pair of approximating samples was not as accurate as for other scenarios. Hence, for this next experiment two pairs of approximating samples ($W = 2$) are used, with one on each of the two peaks in the integrand. The probability of error as a function of noise level is shown in Figure 9.11 with the Monte Carlo results shown in blue, along with the prediction using the Method of Noiseless Integration shown in green and the prediction using the Method of Unbiased Integration shown in black. It can be seen that significant improvement compared to the single pair results of Figure 9.10 has occurred for the Noiseless Integration Method as expected, indicating that it does not do as well extrapolating out to the other mode. On the other hand, the

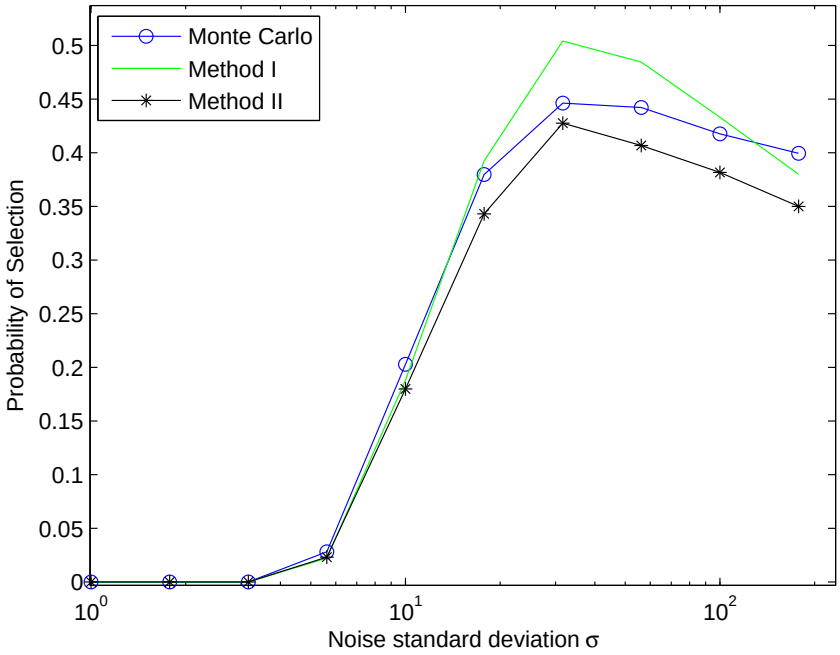

Figure 9.7: Performance prediction for 1-D Gaussian range noise. Probability that the Sentra is chosen as a function of noise standard deviation. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black.

Method of Unbiased Integration shows significant relative improvement for lower noise, but only slight improvement for high noise, which is what might be expected given that it is already a better approximation.

We then run the simulation with four approximating pairs of points ($W = 4$). The two additional pairs are chosen to be symmetric about the primary peak. The results are shown in Figure 9.12. It can be seen that the Unbiased Integration prediction is virtually perfect for this case, and the Noiseless Integration prediction has improved once again. This confirms that the gradual addition of approximating samples will increase the accuracy of the prediction.

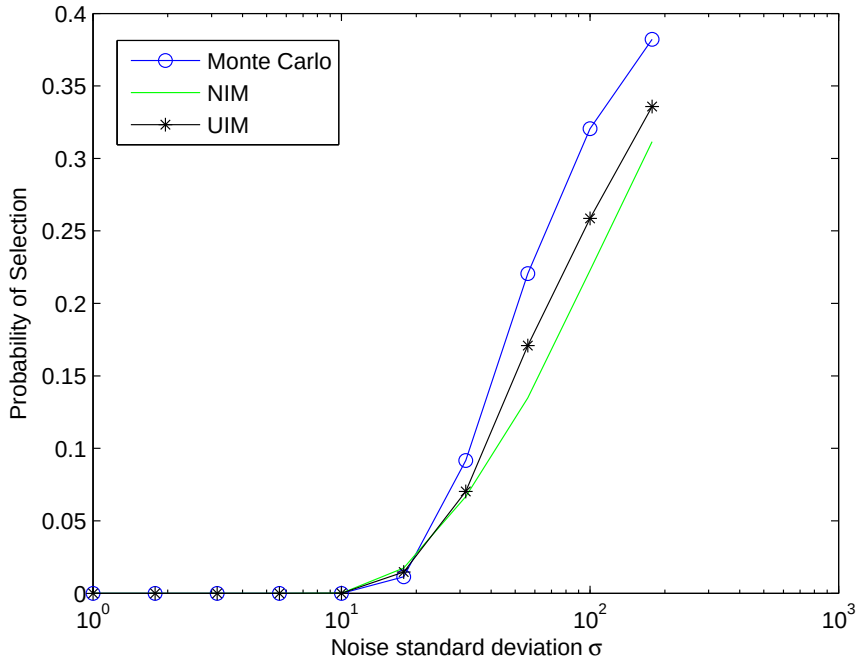

Figure 9.8: Performance prediction for 1-D Gaussian range noise. Probability that the Jeep is chosen as a function of noise standard deviation. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black. Note the initial low probability, which begins to increase as the noise becomes very high.

9.5 Asymptotic Dependence of Performance on Resolution

In this section, we consider the variation of probability of error for a fixed noise variance as a function of sensor resolution. The Gaussian range image noise model is used. Changes in resolution are obtained by downsampling images of size 200×200 pixels. The independent variable is the number of pixels on a side for the downsampled image. The lowest resolution for which we run the experiment is 10×10 pixels.

The results of prediction are shown in Figure 9.13. The Monte Carlo results using 2×10^4 simulations are shown in blue, the Method of Noiseless Integration in green, and the Method of Unbiased Integration in black. The performance gets better rapidly as the image resolution increases, and then saturates out at zero probability of error as expected. As expected, the predictions are accurate, with the UIM outperforming the NIM.

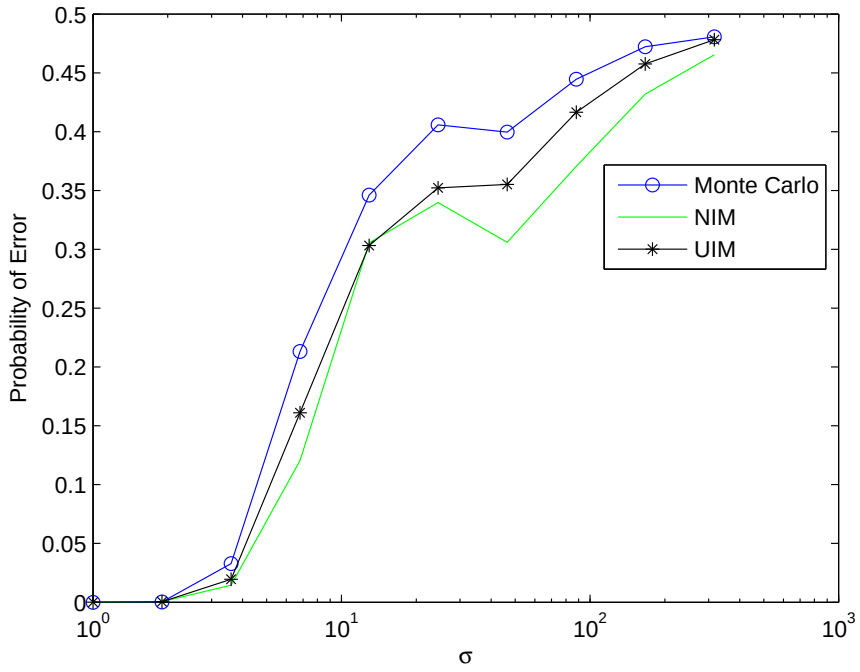

Figure 9.9: Performance prediction for 1-D Gaussian range noise with random deletions with (x, y) position and azimuth angle uncertainty. Probability of error as a function of noise standard deviation shown. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black.

We now use the simple version of the asymptotic approximation for the variation with sensor resolution described in Section 7.1 in (7.10). The results are shown in Figure 9.14, with the Monte Carlo results in blue circles and the Method of Unbiased Integration in black stars. The other six curves are the asymptotically approximating curves based on the Unbiased Integration prediction at each of the six points. As can be seen, the results are quite good in the low error probability region as desired, with the possible exception of the two curves based on the two points with the highest error probability, for which the error is as large as 0.07. The other five curves become overly pessimistic in the high error region, but this should not be an issue since it is not desired to operate in that region anyway. Otherwise, the error is within 0.03. Since this asymptotic prediction requires merely the evaluation of a series of Q functions to generate the resolution curve, once one point has been predicted, the additional error incurred with this method is outweighed by its very large increase in speed.

Finally, we consider the more complex asymptotic prediction method that also requires the

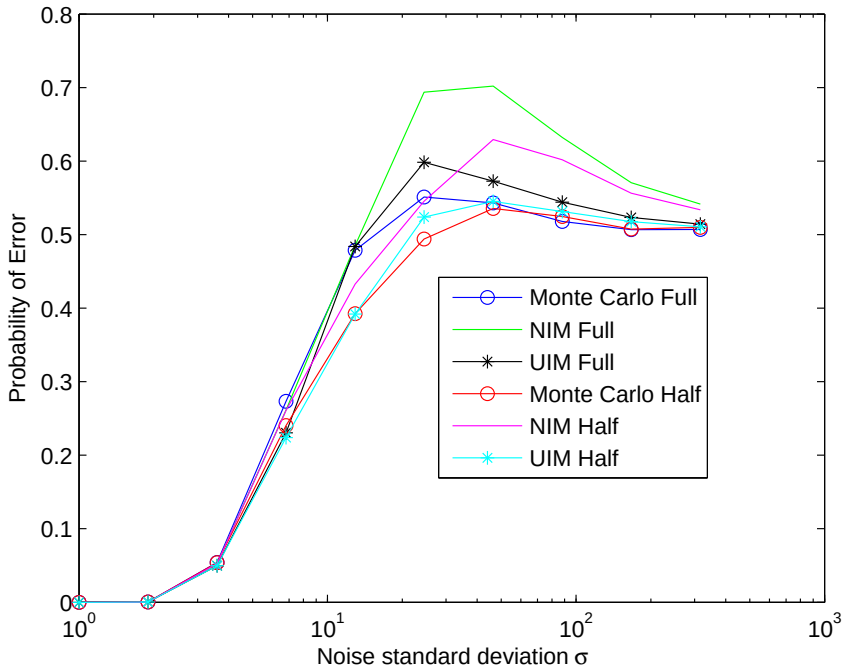

Figure 9.10: Performance prediction for Gaussian range image noise with uniform anomalies. Nuisance parameter is target angle with prior $-\pi$ to π. Probability of error given H_0 is true as a function of Gaussian parameter σ. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black. Results for half prior of $-\pi/2$ to $\pi/2$ also shown, with Monte Carlo simulation results shown in red, NIM shown in magenta, and UIM shown in cyan. Note the degradation in performance in moving to the more realistic prior.

threshold adjustment at the point to be predicted (7.9). As can be seen, the results are in general significantly more accurate, although there are points where the approximation is slightly worse. All the curves except one stay close to the Monte Carlo truth for all noise levels, unlike for the previous method. The one exception to this is the curve based on the point with the highest error probability, which has an overly high error prediction for almost all of the range. This is because the threshold adjustment at that point is dramatically smaller than for the other points, since at that point, the curve is beginning to saturate out near 0.5 error, and the integrands are becoming relatively uniform over the angle space. For extrapolation using this point, the simpler asymptotic approximation is more reasonable.

It thus appears that the asymptotic techniques for resolution variation developed in this thesis provide highly efficient and reasonably accurate methods of predicting the performance at a certain

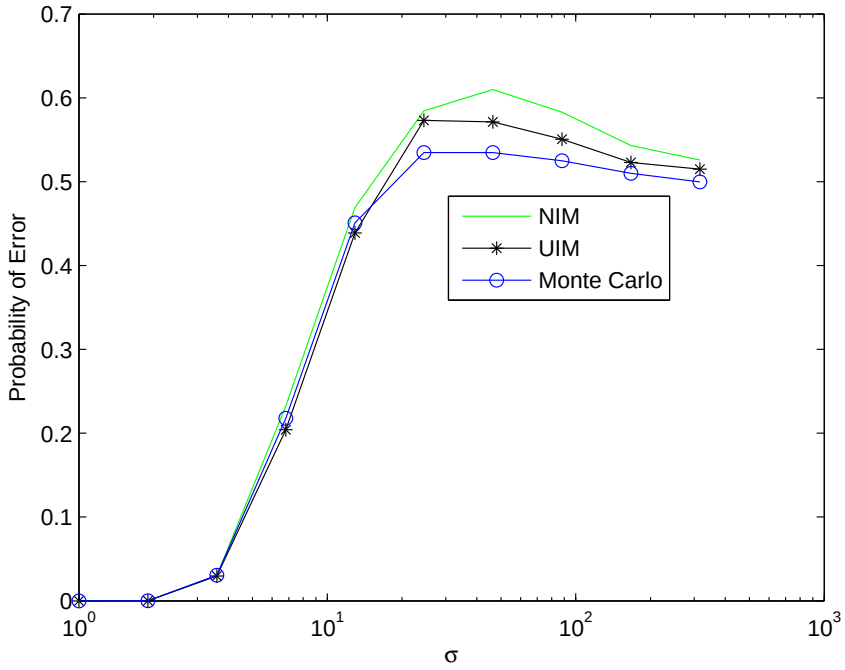

Figure 9.11: Performance prediction for Gaussian range image noise with uniform anomalies. Nuisance parameter is target angle with prior $-\pi$ to π, and two pairs of approximating samples are used to predict performance. Probability of error given H_0 is true as a function of Gaussian parameter σ. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

image resolution given a predicted performance at another resolution and the same noise level. This also confirms that even fairly small images (25 pixels on a side) have high enough resolution to apply these asymptotic techniques. This is further supported by the fact that it is not limited to any one noise model, so long as the Central Limit Theorem applies.

9.6 Dependence of Performance on Anomaly Probability

In this section, we consider the application of the results for the variation of performance with probability of anomaly. We use the methods of Section 6.1.3.

We again use the Honda Civic and Toyota Avalon models as targets, with the nuisance parameter being target azimuth angle with a uniform prior from $-\pi$ to π, and the image size 56×56. We use the Gaussian noise with random deletion model, with the noise variance held constant as

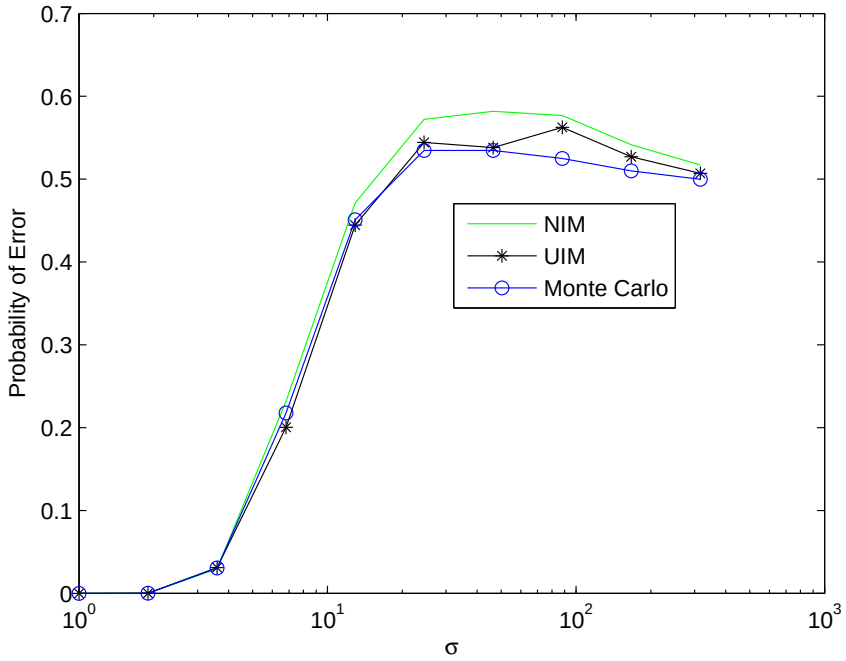

Figure 9.12: Performance prediction for Gaussian range image noise with uniform anomalies. Nuisance parameter is target angle with prior $-\pi$ to π, and four pairs of approximating samples are used to predict performance. Probability of error given H_0 is true as a function of Gaussian parameter σ. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

the probability of anomaly varies from 0 to 0.7. The results for the predictions of the Noiseless Integration Method and the Method of Unbiased Integration are shown in Figure 9.16, with the independent variable being the probability of anomaly. The Monte Carlo results are shown in blue and are generated using 2×10^4 simulations, the results of the Noiseless Integration Method are shown in green, and the results of the Method of Unbiased integration are shown in black. For some of the range of values, the Noiseless Integration prediction is slightly better. This is likely a coincidence resulting from the full prior and the inherent target symmetries, as it appears that this is biasing the UIM prediction lower, and the NIM tends to have a higher error prediction than the Unbiased Integration prediction, thus accidentally correcting for the Unbiased Integration predictions errors. This is confirmed by the divergence of the Noiseless Integration prediction for high error probabilities.

To test the results of this approximation for other anomaly models, we repeat this experiment using the coherent LADAR model. The noise variance is held constant as the probability of anomaly

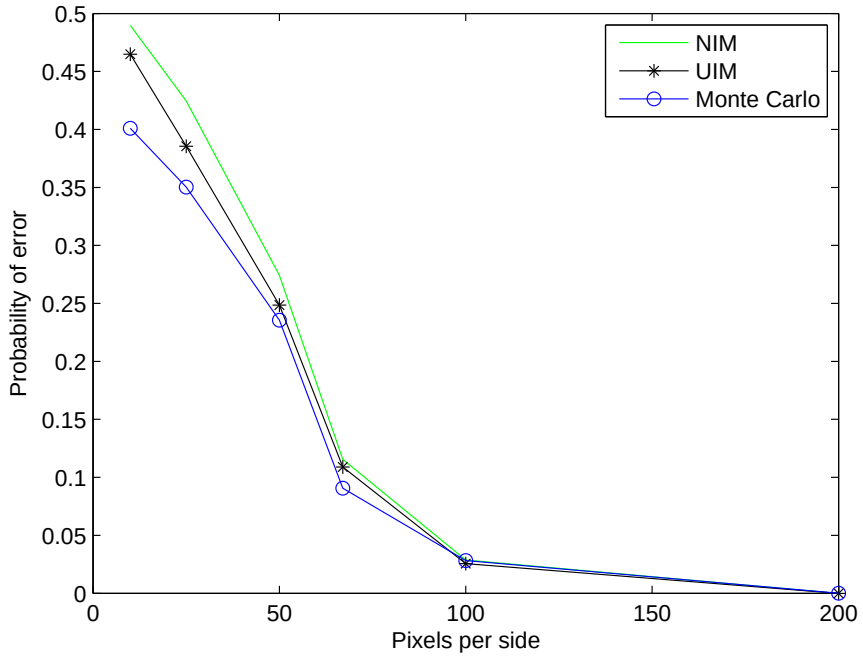

Figure 9.13: Performance prediction for 1-D Gaussian range noise. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

varies from 0 to 0.7.

The approximation for the variation of error rate with probability of anomaly given in Section 6.1.3 in (7.11) is then implemented. To test this method, we use the coherent LADAR model with the uniform anomaly distribution. The probability of anomaly is varied by varying the length of the uniform distribution in such a way as to preserve the anomaly density. The width of the uniform distribution is chosen to be 20 times the Gaussian standard deviation for the lowest nonzero probability of anomaly. The only exception to this procedure for determining the uniform distribution length is for the case of zero probability of anomaly, for which there is no anomaly component. For the reference point needed to apply (7.11), we use the Method of Unbiased Integration, and generate 5 curves, each based on the reference point at a different odd-numbered point on the complete prediction curve, with the count starting at zero probability of anomaly. The results are shown in Figure 9.17. The Unbiased Integration prediction is shown in black, and the five remaining curves are the simple approximation curves generated using the reference point indicated, with the num-

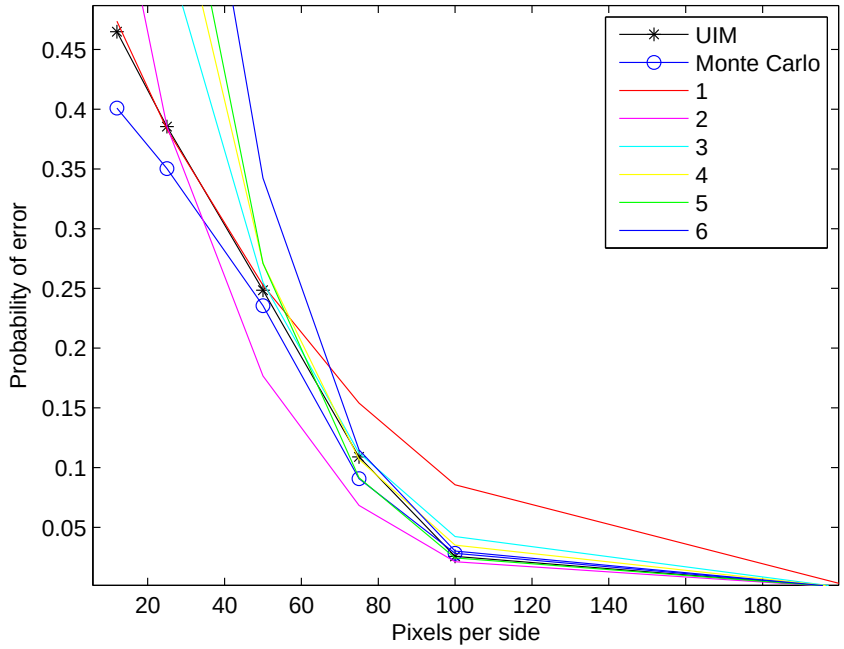

Figure 9.14: Performance prediction for 1-D Gaussian range noise. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue, UIM shown in black. Other six curves use the simple resolution asymptotic method and are numbered by the point used to generate the curve, with the points numbered from left to right.

bering being from left to right. It can be seen that the simple approximation is reasonably accurate, given the speed of computation. This confirms the usefulness of the simple anomaly rate variation approximation.

9.7 Performance as a Function of Sensor Elevation Angle

We now perform an experiment to find the variation of the probability of error with sensor elevation angle. The Gaussian range noise with random deletions model is used, with a probability of anomaly 0.2. The Civic and Avalon are used as the targets, and the elevation angle is varied from 10 degrees to 90 degrees in steps of 10 degrees, with the azimuth angle and noise variance held constant. The image size is again 56×56 pixels. The true target was the Civic in a pose at a 45 degree angle to the sensor line of sight, and the nuisance parameter was the target azimuth angle. The prior is uniform

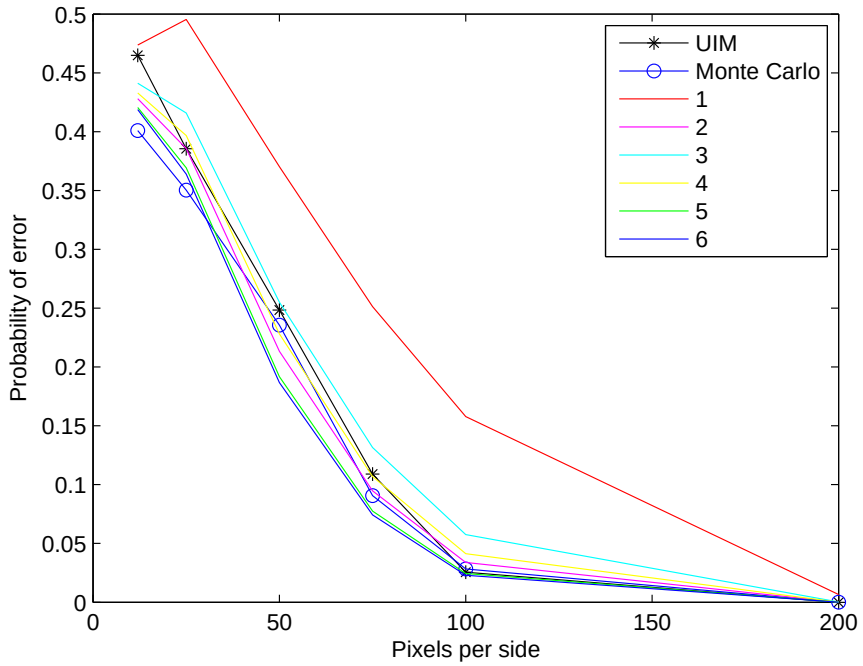

Figure 9.15: Performance prediction for 1-D Gaussian range noise. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue, UIM shown in black. The other six curves use the simple resolution asymptotic method and are numbered by the point used to generate the curve, with the points numbered from left to right.

from $-\pi$ to π.

The probability of error results are shown in Figure 9.18. The Monte Carlo results using 2×10^4 simulations are shown in blue, and the predictions using the Noiseless Integration method and Unbiased Integration method are shown in green and black respectively. It can be seen that the performance is very good for low elevation angles and degrades quickly to about 0.25 probability of error around 50 degrees elevation before leveling out and slightly improving. This degradation in performance is to be expected, since low elevation angles result in a higher contrast between the target and its flat ground plane background, whereas for high elevation angles, the difference in range between the target and the background is only as large as the target height. Since the difference in target outline is very important to the classifier, high contrast with the background, as it improves the ability to segment the target, improves the classification performance. Due to the nature of elevation angle, the rate of improvement in contrast quickly slows and practically stops as

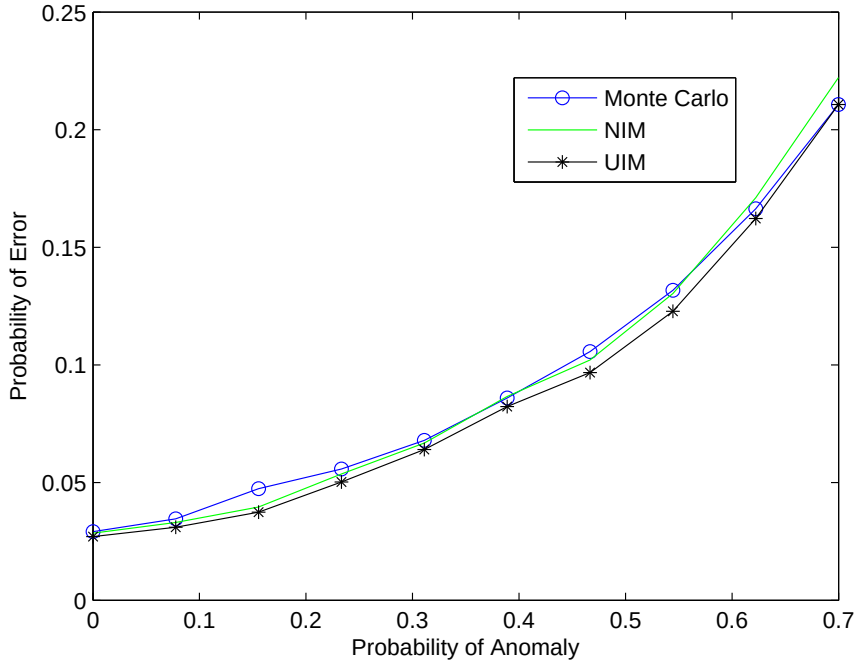

Figure 9.16: Performance prediction for 1-D Gaussian range noise with random deletions. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue, NIM shown in green, and UIM shown in black.

the elevation angle increases above about 45 degrees, which is what we observe in the results.

The simple approximate elevation angle variation results from Section 7.2.3 for low elevation angles are then applied to this experiment. The assumption on which this is based is that the number of pixels on target does not change significantly as the elevation angle varies. Four approximate curves are generated using as reference samples the Unbiased Integration predictions for elevation angles of 20, 40, 60, and 80 degrees, with each reference sample uniquely associated with one of the approximating curves. The results are shown in Figure 9.19. The Monte Carlo results and the Unbiased Integration prediction are shown for clarity. The approximating curves are the remaining four curves. Considering the simplicity of the approximation, the results are quite accurate. The approximation results are very accurate for the region before the error probability levels out, but degrade for the leveled out region where the probability of error is high as expected based on the assumption of low to medium elevation angle. The predictions generated using reference points in the leveled out region are more accurate than those generated farther away. This result indicates

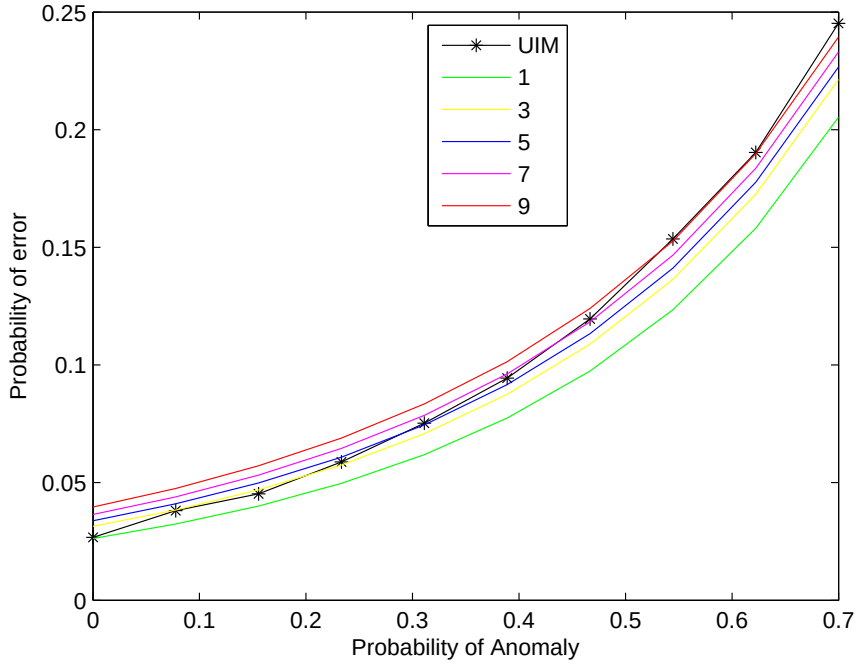

Figure 9.17: Performance prediction for 1-D Gaussian range noise with uniform anomalies. Probability of error given H_0 is true as a function of image dimension in pixels. UIM prediction shown in black. The five remaining curves are the predictions using the simple method of prediction, using different points as reference points.

that this approximation to the performance variation is a reasonably accurate method of rapidly obtaining a prediction for the performance variation with elevation angle.

The high elevation angle approximate results from Section 7.2.3 are then applied to the same data. The same reference points are used, and the results are shown in Figure 9.20. As expected, the approximations for the high elevation angle region using reference points also in that region are significantly more accurate than the was the case for the low elevation angle approximation. Also as expected, the approximation using the low elevation angle of 20 degrees as a reference is quite poor, in contrast to its high degree of low angle accuracy for the low angle approximation. For the mid level elevations, both approximate prediction methods seem to be reasonably accurate and in agreement. From this (if a few samples are available) it can be determined approximately what the cutoff is between "low" and "high" elevation angles. It should be noted, though, that as most of the variation in performance occurs in the low angle range, this is likely to be the most interesting and critical area, as opposed to the relatively constant high elevation angle region.

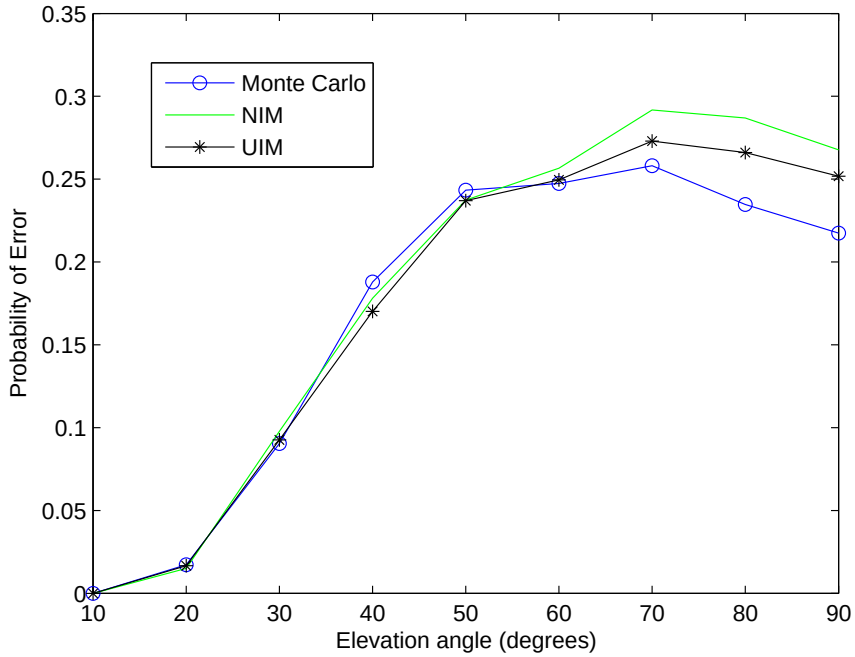

Figure 9.18: Performance as a function of sensor elevation angle in degrees. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue, NIM in green, and UIM shown in black.

9.8 Performance as a function of Angular Pose

We consider the variation of performance with respect to the true value of the target azimuth angle. We use the targets of Figures 9.1 and 9.2, with the former being the true target, and the elevation angle being 40 degrees. The image is 56×56 pixels in size. The Gaussian noise plus random deletions noise model is used, with probability of anomaly 0.2. We use the Method of Unbiased Integration to predict performance, calculating the LLR pdf using the Gaussian pdf approximation since the number of pixels is large enough for the Central Limit Theorem to apply. The azimuth angle prior is uniform from $-\pi$ to π radians. We vary the azimuth angle of the true target in steps of 0.2 degrees from 0 to 360 degrees and predict the probability of error for each. The results are shown in Figure 9.21.

As can be seen, the probability of error varies significantly with pose, from about 0.2 to 0.4. A view of the true target in the pose with maximum performance is shown in Figure 9.22, and a view in the pose with minimum performance in Figure 9.23. As can be seen, the number of pixels on

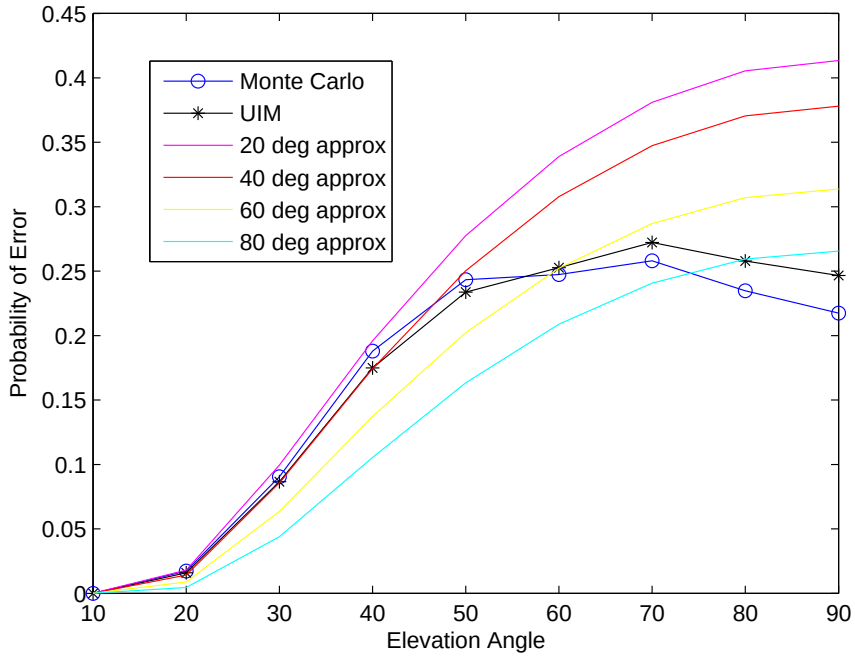

Figure 9.19: Performance as a function of sensor elevation angle in degrees. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue and the Unbiased Integration prediction is shown in black. Remaining four curves are the simple approximation curves generated using the indicated reference points.

target is significantly larger for the image in the pose with maximum performance than in the image in the pose with minimum performance.

A significant portion of this effect is due to variations in the number of pixels on target. First, we show a plot of the number of pixels on target as a function of pose angle in Figure 9.24, showing that the number of pixels on target varies widely, from about 600 to 800. We then use the asymptotic prediction for the variation of performance with resolution to generate a plot of the asymptotic prediction of what the performance curve would be if the resolution was altered for each plot such that the image always had the same number of pixels on the target, equal to the average number. The results are shown in Figure 9.25, with the true performance shown in blue and the performance normalized to the same number of pixels shown in red. Note that the red curve is flatter than the blue, but still varies considerably, indicating that a large portion the variation in performance with azimuth is not due to changes in the number of pixels on target. The curve appears to be almost

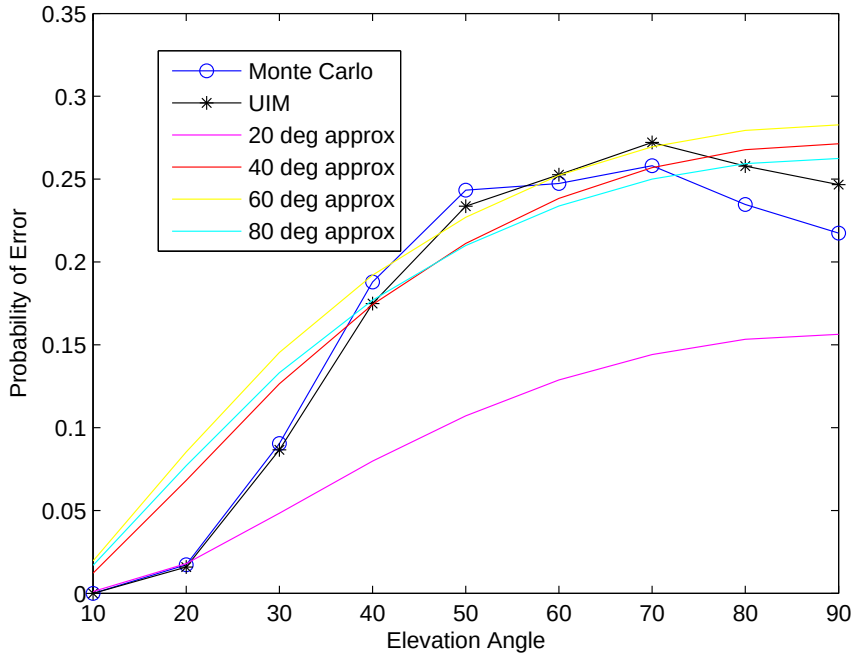

Figure 9.20: Performance as a function of sensor elevation angle in degrees. Probability of error given H_0 is true as a function of image dimension in pixels. Monte Carlo simulation results shown in blue and the Unbiased Integration prediction is shown in black. Remaining four curves are the simple approximation curves generated using the indicated reference points.

piecewise constant, however, indicating that the targets are easier to separate when looking at the front of the vehicle than when looking on the back, even on a per pixel basis. To confirm this, we show a plot (Figure 9.26 of the expectation of the pseudo-loglikelihood ratio $\log(L')$ divided by the number of pixels on target, where it appears that the average separation of pixels is smaller for the region for which the probability of error is higher. Some of this variation may be reduced by adding in target position uncertainty to the azimuth uncertainty, but overall, from this result, it appears that the performance of ATR algorithms should vary with pose, or they will be suboptimal.

An additional result from this experiment is the unconditional probability of error. That is, the predictor does not know the true pose of the target, only its prior distribution. To obtain the unconditional performance, the true pose must be integrated out. That is,

$$P(e|H_0) = \int_S P(e|H_0, \underline{\theta}) f_0(\underline{\theta}) d\underline{\theta} \tag{9.1}$$

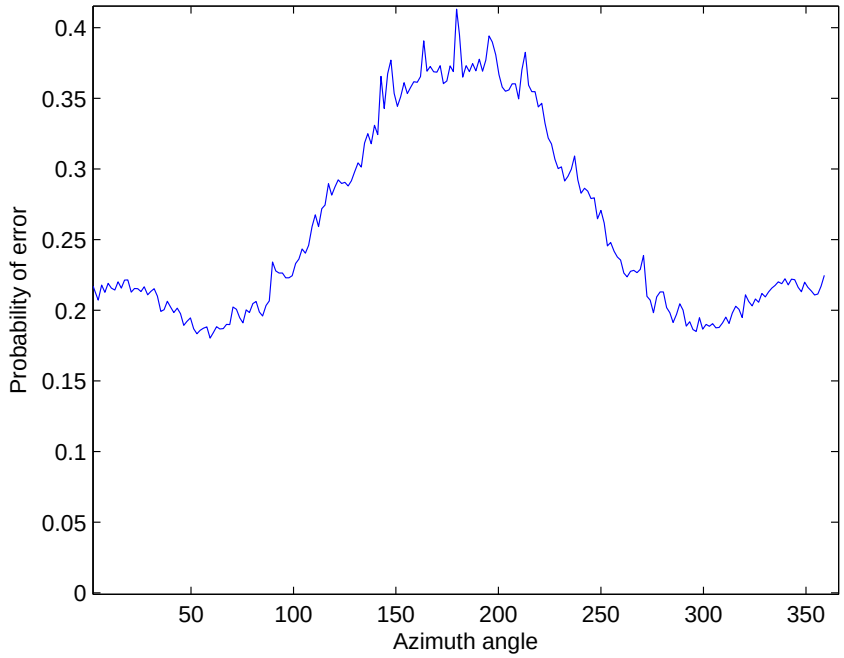

Figure 9.21: Performance prediction for 1-D Gaussian range noise with random deletions. Probability of error given H_0 is true as a function of truth azimuth angle in degrees. Prediction using Method of Unbiased Integration shown.

where f_0 is the prior on the true target nuisance parameters as before.

The results obtained in this subsection were derived using a uniform prior, so the resulting unconditional probability of error is 0.26. This method can be used to find the change in performance resulting from implementing a Bayesian classifier using the incorrect prior, by using different priors to predict conditional performance and to do the integration to obtain the unconditional performance. This is then compared to the predicted performance using the true prior.

9.9 Pixel Contributions to Performance

In this section, we examine which pixels in an image are most important to the separability of a pair of targets. We consider a scenario where the true target is the Honda Civic in the pose shown in Figure 9.27. The image size is 56×56 pixels. The alternative target is the Toyota Avalon in the same pose. The nuisance parameter is azimuth angle. The Gaussian noise plus random deletions model

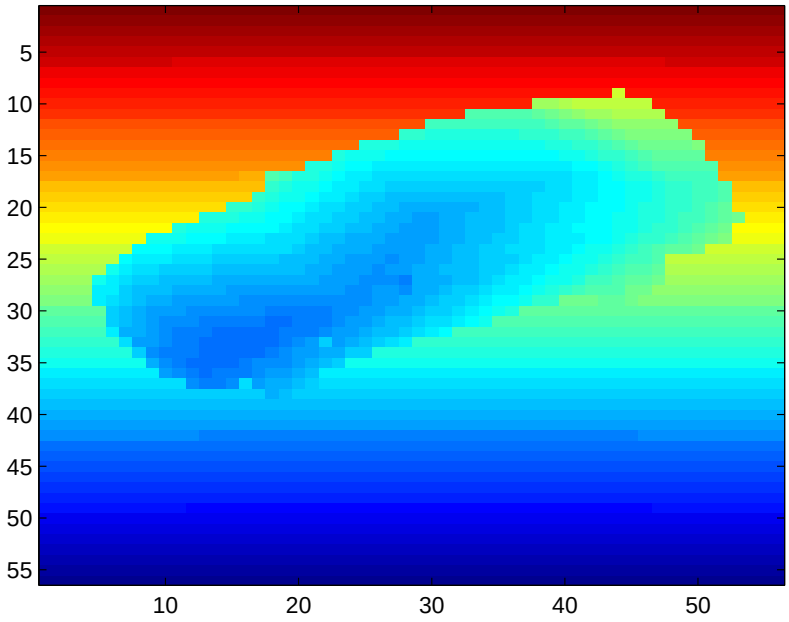

Figure 9.22: View of true target (Civic) in the pose for which the performance is best. Note the large number of pixels on target.

was used with a probability of deletion 0.2. The noise level is chosen such that the probability of correct classification is about 0.75.

We consider two measures of the usefulness of each pixel to the recognition task. Both are based on Grenander's method of predicting performance. For this method, the performance prediction is based on two quantities, the pseudo loglikelihood ratio L' (4.26) and the threshold adjustment ν' (4.13), which is the square root of the ratio of the second derivatives of the loglikelihood functions for the two targets, evaluated at a certain value of the nuisance parameter. Since loglikelihood is additive on a per pixel basis due to independence (4.26), the two measures we consider are the expected value of each pixel's contribution $E[L'(s_k)]$ to the pseudo loglikelihood ratio, and the ratio of the second derivatives $\ddot{E}_i(s_k, H_i, \theta)|_{\theta_i^*}$ of each target's loglikelihood of the pixel of interest, evaluated at the same point the threshold adjustment is evaluated. The first measure, by definition, measures pixel contributions to the minimum target separability over the nuisance parameter space, and is strictly additive. The second measure measures how strongly a given pixel moves the threshold adjustment, and in which direction. This latter measure is not additive, but gives a

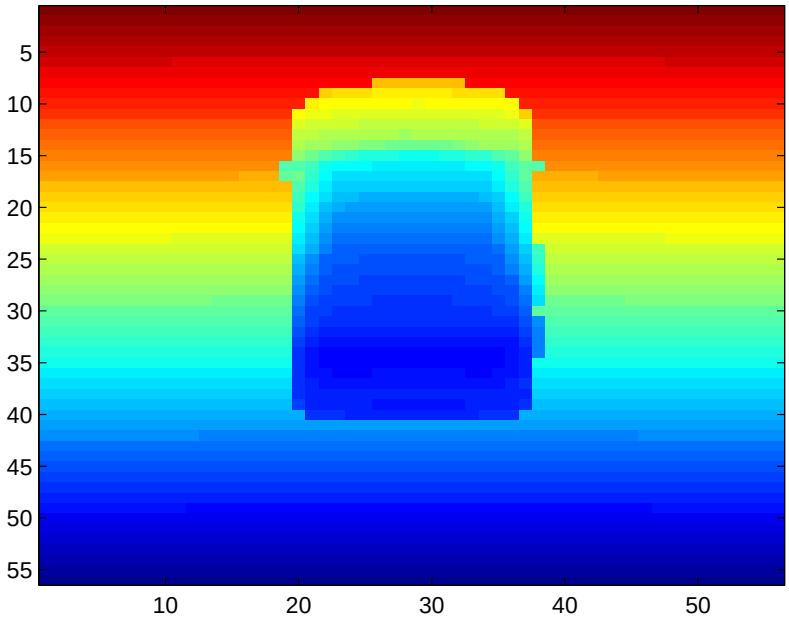

Figure 9.23: View of true target (Civic) in the pose for which the performance is worst. Note the relatively small number of pixels on target.

general indication of which pixels are the most useful in determining the target's exact pose and the resulting effect on the performance.

A plot of the first measure is shown in Figure 9.28. Positive values indicate that the expected value of the pseudo loglikelihood ratio is biased towards the true target, with the magnitude indicating the strength of that bias. It can be seen that the most important portions of the image for recognition are along the edges of the vehicles, which is as expected, since it is there that the differences in outline occur. The most important edge regions appear to be the front and back, along with the sideview mirror. By far the most significant region is the rear bumper area, where the viewing angle results in a significant difference in target outline. In this region the differences are made particularly large, partially because only the top part of the rear of the vehicle can be seen, and because the low viewing angle amplifies the discontinuity between the targets trunk top and the ground.

A plot of the second measure is shown in Figure 9.29. The results were obtained using 5 point finite differencing to obtain the second derivatives. The ratio of the second derivatives is the ratio of that for the alternative target to that of the true target. It can be seen that the pixels that have

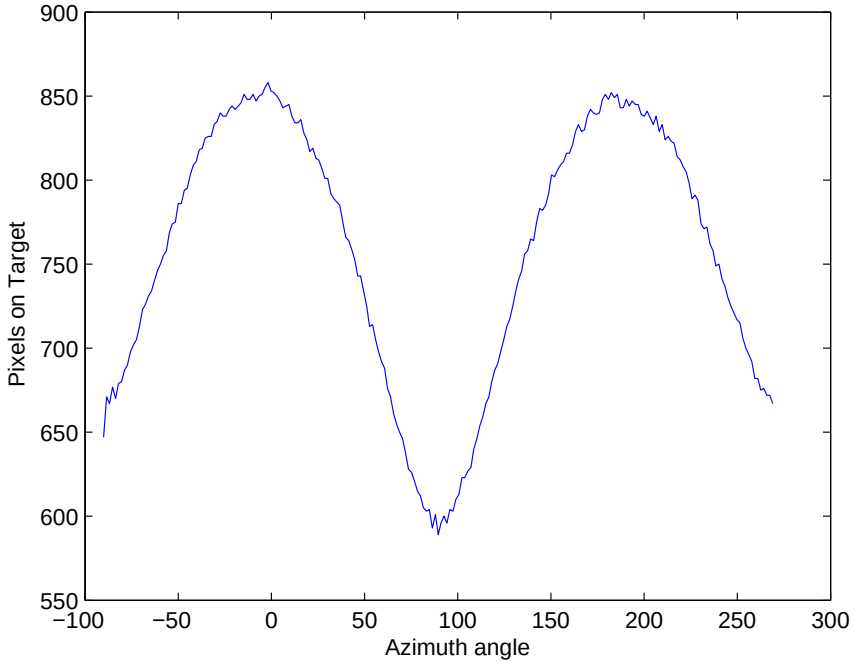

Figure 9.24: Number of pixels on target as function of Honda Civic azimuth pose angle in degrees. Note that the minima and maxima correspond approximately to the maxima and minima of the performance plot.

the most effect on the threshold adjustment are related to the side view mirrors, the front end of the vehicle, and the trunk area. This is as expected.

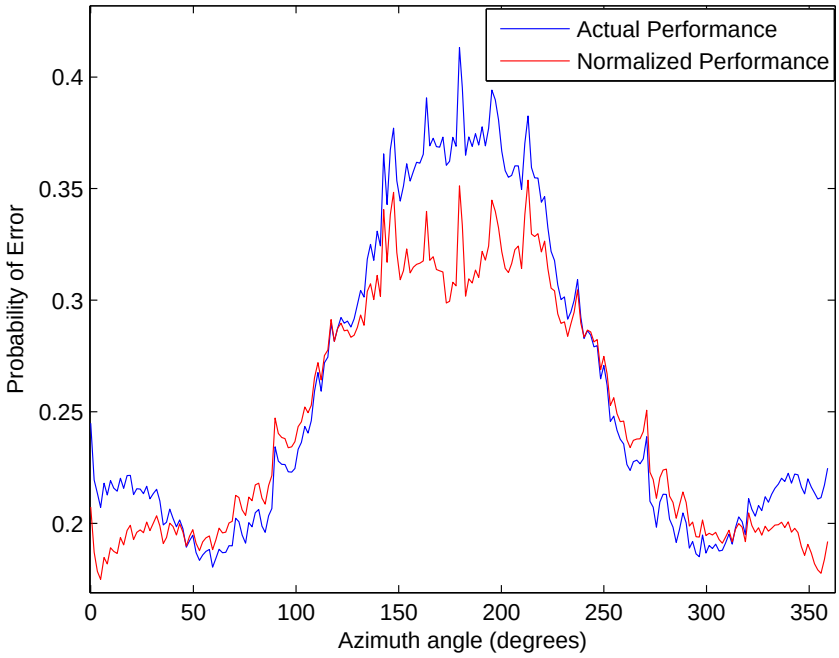

Figure 9.25: Performance prediction for 1-D Gaussian range noise with random deletions. Probability of error given H_0 is true as a function of truth azimuth angle in degrees shown in blue. The probability corrected for the number of pixels on target is shown in red. Prediction using Method of Unbiased Integration shown.

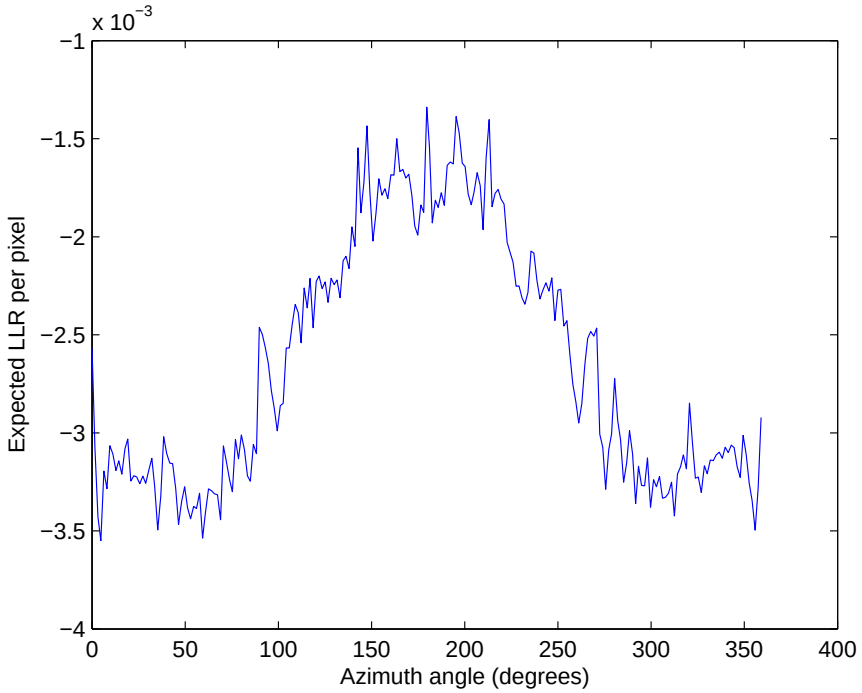

Figure 9.26: Plot of the expected pseudo LLR divided by the number of pixels on target as a function of azimuth angle.

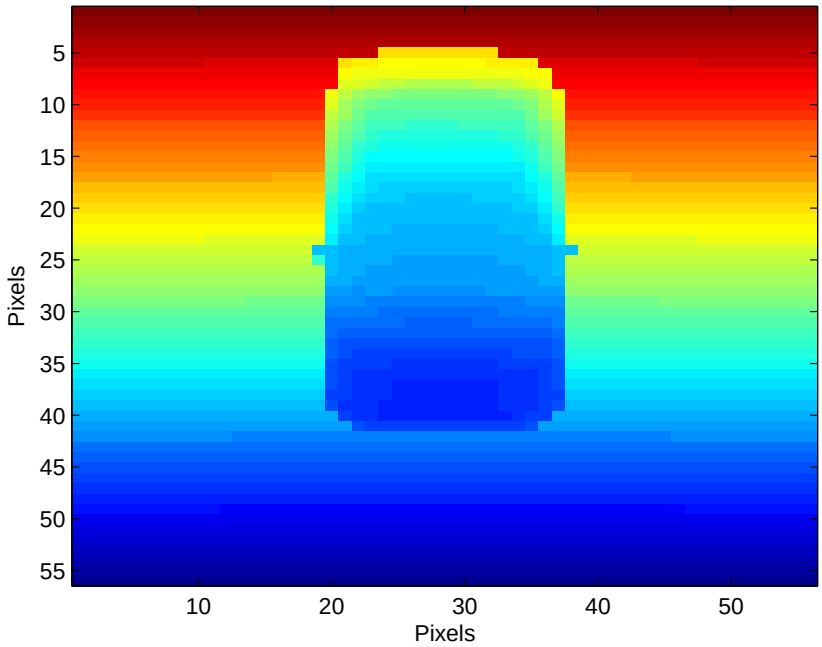

Figure 9.27: View of Civic in pose used for individual pixel analysis.

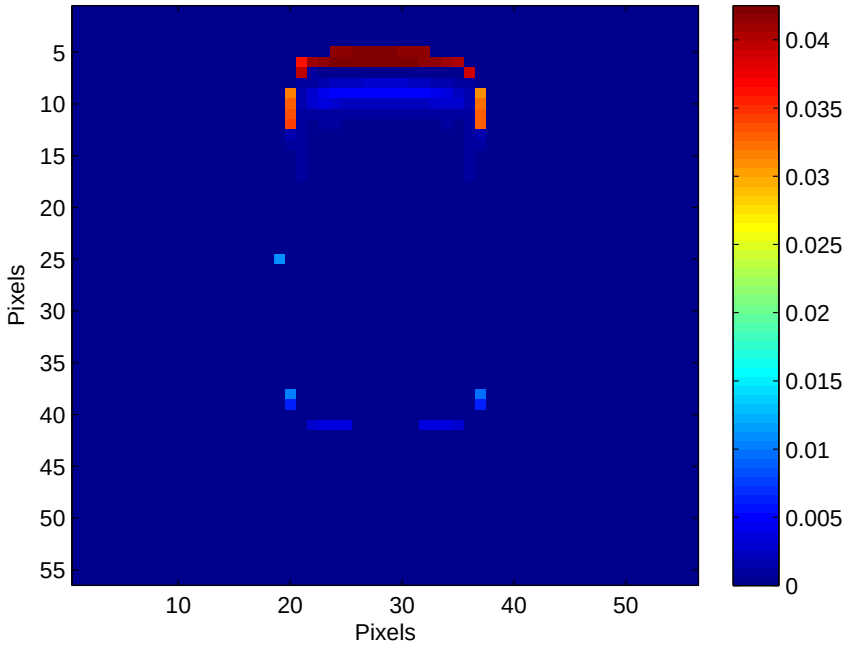

Figure 9.28: Expected value of pseudo loglikelihood ratio. Positive values indicate bias towards true target. Note the high magnitude points along outline difference at rear bumper.

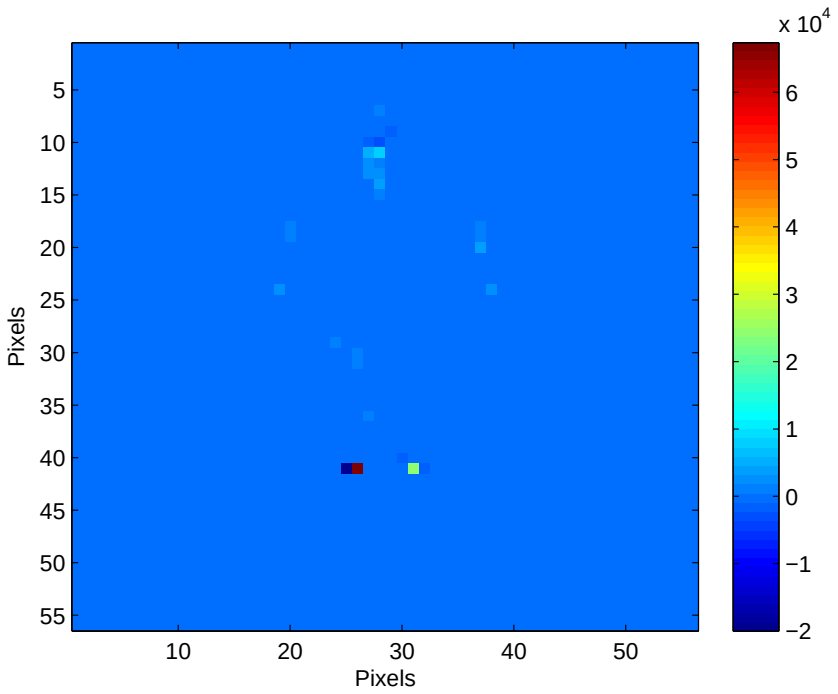

Figure 9.29: Pixel contributions to the Laplace approximation threshold adjustment. Second derivative of loglikelihood function of Avalon target divided by that of the true target.

Chapter 10

Conclusion

Two new methods of predicting the performance of the Bayesian ATR classifier were presented, the Noiseless Integration Method, and the Unbiased Integration Method. Both methods involve the integration of two deterministic scalars across the nuisance parameter space and using the result to adjust the threshold on a pseudo log likelihood ratio test. The first method is a simple extension of the method of Grenander giving significantly increased prediction accuracy. The second method is more accurate than the first as it is based on an unbiased approximation of the Bayesian integral. The new methods were then compared with the method of Grenander in an experiment, and both of these methods outperformed the method of Grenander by a large margin. The improvement in performance over the method of Grenander was found to be especially significant when significant target near symmetry was present, and when noise levels are medium to high. This is largely because the method of Grenander assumes that no target symmetry is present and because the Laplace approximation it is based on is only reasonably accurate for low noise. This increase in accuracy comes at an increase in computational cost. For any prediction method, it is essential that the computational intensity of the prediction is much less than that of a similarly accurate Monte Carlo simulation. Each of the methods presented here can easily achieve this requirement, as the complexity is slightly more than that of a single Bayesian recognition operation.

These new methods were then applied to several LADAR noise models of varying complexity and accuracy, allowing for significantly more accuracy than provided by the simple Gaussian mod-

els used previously for LADAR performance prediction. Modeling capabilities added using these models include the ability to model anomalous pixels and cross-range noise, as well as being able to handle point clouds. From experiments, it was found that the classification performance can be significantly optimistic if an overly simplistic noise model is used for prediction, thus necessitating the use of accurate noise models to accurately predict performance. To allow the derivation of more analytic results, a random deletion pixel anomaly model was also used as a successful approximation to the more accurate anomaly models.

Various extensions to these two methods were developed. A method of predicting performance for a multiple target classification problem was developed, using the joint distribution of M log likelihoods and the Bayesian MAP classifier. A method of gradually increasing the accuracy of the prediction at the cost of increasing complexity is also developed. This method provides a smooth transition from the simplest version of the two methods introduced here and a complete Bayesian Monte Carlo simulation. Variations of the two methods were developed for predicting the performance of classifiers derived using incorrect noise models.

Simple analytic asymptotic expressions were then developed for the variation of the probability of error with sensor resolution, probability of anomaly, sensor elevation angle, and noise variance. Most of these methods require the computation of a more accurate prediction for one point along the curve before the entire curve is generated using these methods. As a result, they are significantly more accurate than is possible using the same simplifying assumptions but without the initial point. These approximations were found to apply to a large range of the relevant parameter values, and to be quite accurate for all reasonable scenarios tested.

Experiments were then run with several different target CAD models and different noise models to generate performance curves as functions of different parameters of interest. It was found that both the methods presented here perform quite well, especially the Method of Unbiased Integration. Comparing the NIM and the UIM, the UIM is clearly the better prediction method, giving much more accurate and more consistent results for high noise, being able to predict performance for discontinuous noise pdfs, and showing significantly better accuracy for the case of incorrect classifier noise models. Which of these methods are selected in a given situation will depend on how

much accuracy is required and whether or not the expectations required by the Method of Unbiased Integration are analytically tractable.

Several additional observations were discussed. First, it was found that the effect of the threshold adjustment ν' is relatively small for very low noise, but becomes significant for high noise. It was also found that for noise levels that are not very small the degree of rotational near symmetry commonly found in common civilian targets is sufficient to cause the Laplace approximation to result in significantly inaccurate predictions, as it assumes no near symmetries are present. The use of numerical integration causes the presence of near symmetry to affect NIM and UIM to a much lesser degree. Finally, it was observed that accuracy in performance prediction depends heavily on the accuracy of the noise model used. Especially for low noise, accuracy in noise modeling affects prediction accuracy much more than does the choice of prediction method.

In conclusion, we have developed a set of improved techniques of varying complexity and accuracy to predict the performance of LADAR ATR given CAD models of the relevant targets. These methods were implemented for a variety of noise models not previously used for performance prediction, some of which allow for significantly higher accuracy. The prediction results were confirmed experimentally using Monte Carlo simulations.

Appendix A

Proof of Theorem I

In this appendix, we prove the following theorem. The approximation in (4.22) is asymptotically correct, that is

$$\int_S p(\mathcal{I}|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta} \sim p(\mathcal{I}|H_i,\underline{\theta}_i^*)\frac{\int_S E[p(\mathcal{I}|H_i,\underline{\theta})]f_i(\underline{\theta})d\underline{\theta}}{E[p(\mathcal{I}|H_i,\underline{\theta}_i^*)]}$$

As the noise level goes to zero the expectation of a function converges to the value of the function when the noise level is zero. This is because if a random variable $x(\sigma)$ with variance σ^2 has an expectation $E[x(\sigma)]$ that is continuous with respect to σ

$$\lim_{\sigma\to 0}\frac{E[x(\sigma)]}{x(0)} = \frac{E[x(0)]}{x(0)} = 1 \tag{A.1}$$

As a result,

$$\int_S E[p(\mathcal{I}|H_i,\underline{\theta})]f_i(\underline{\theta})d\underline{\theta} \sim \int_S p(\overline{\mathcal{I}}_0^*|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta} \tag{A.2}$$

$$E[p(\mathcal{I}|H_i,\underline{\theta}_i^*)] \sim p(\overline{\mathcal{I}}_0^*|H_i,\underline{\theta}_i^*)$$

which results in

$$\frac{\int_S E[p(\mathcal{I}|H_i,\underline{\theta})]f_i(\underline{\theta})d\underline{\theta}}{E[p(\mathcal{I}|H_i,\underline{\theta}_i^*)]} \sim \frac{\int_S p(\overline{\mathcal{I}}_0^*|H_i,\underline{\theta})f_i(\underline{\theta})d\underline{\theta}}{p(\overline{\mathcal{I}}_0^*|H_i,\underline{\theta}_i^*)} \tag{A.3}$$

by the definition of asymptotic equality and the properties of the ratio of limits.

Thus it is apparent that the Noiseless Integration Method and the Unbiased Integration Method are asymptotically equivalent, since the right hand side of the equation is identical to the expression in (4.14). Since the Noiseless Integration Method has been shown to provide an approximation that is asymptotically equivalent to the truth [1], the Unbiased Integration Method does also.

Appendix B

Asymptoticity of Deletion Approximation for Coherent LADAR Anomalies

In this appendix, it is shown that the Gaussian noise with random deletions model in Section 3.2.3 is asymptotically equivalent to the coherent LADAR model in Section 6.2.1 as the range swath $b - a \to \infty$ for constant probability of anomaly α. As the loglikelihood ratio is the quantity of interest, we consider the difference in $E[\log p(s_k|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*]$ for both of the models. This is the expectation of each pixel's contribution to the loglikelihood ratio, split between the two hypotheses. Let the difference in the expectations be

$$\delta = E_{coh}[\log p_{coh}(s_k|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*] - E_{del}[\log p_{del}(s_k|H_i, \underline{\theta})|H_0, \underline{\theta}_0^*] \tag{B.1}$$

where the subscripts coh and del indicate the coherent LADAR and the random deletion models respectively. The definitions for each noise model are given in the sections listed above. Let the coherent LADAR range swath be $[a, b]$. Let the Gaussian noise standard deviation be σ, the mean $\overline{s}_{k,i}$ and the probability of anomaly α_i for both models and hypotheses. For the Gaussian noise with random deletions model, let the anomaly subset $\mathcal{B}$ be such that $|\mathcal{B}| = b - a$ for accuracy in comparison.

We have

$$\delta = \int_a^b \log(p_{coh}(s_k|H_i,\underline{\theta}))\left(\frac{1-\alpha_0}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2\}+\frac{P_a^{(0)}}{b-a}\right)ds_k \tag{B.2}$$
$$-\int_a^b \frac{\alpha_0}{|\mathcal{B}|}\log\frac{P_a^{(i)}}{|\mathcal{B}|}ds_k$$
$$-\frac{(1-\alpha_0)}{\sqrt{2\pi\sigma^2}}\int_a^b \log\left(\frac{(1-\alpha_i)}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2\}\right)\exp\left\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2\right\}ds_k$$

Separating this into two portions

$$\delta_1 = \int_a^b \log\left(\frac{1-\alpha_i}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_{k,i})^2\right\}+\frac{\alpha_0}{b-a}\right)\frac{\alpha_0}{b-a}ds_k \tag{B.3}$$
$$-\int_a^b \frac{\alpha_0}{|\mathcal{B}|}(\log\frac{\alpha_i}{|\mathcal{B}|}ds_k$$
$$\delta_2 =$$
$$\int_a^b \log\left(\frac{1-\alpha_i}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(s_k-\overline{s}_{k,i})^2}+\frac{\alpha_0}{b-a}\right)\left(\frac{1-\alpha_0}{\sqrt{2\pi\sigma^2}}\exp\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2\}\right)ds_k$$
$$-\frac{(1-\alpha_0)}{\sqrt{2\pi\sigma^2}}\int_a^b \log\left(\frac{(1-\alpha_i)}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2}\right)\exp\left\{-\frac{1}{2\sigma^2}(s_k-\overline{s}_k^{(0)})^2\right\}ds_k$$

Since the coherent loglikelihood in the region outside the area of the Gaussian peak is asymptotically equal to $1/\log(b-a)$ due to the rapid decay of the Gaussian portion of the likelihood, and since the smaller likelihoods have the highest magnitude loglikelihoods and thus dominate the expectation, and because the width of the Gaussian distribution is fixed by σ, we have that

$$\lim_{b-a\to\infty}\delta_1 = 0 \tag{B.4}$$

For δ_2, it also is apparent that the expressions within the logarithms converge as $b-a$ increases. Since the remainder of the integrands is the same, we have that

$$\lim_{b-a\to\infty}\delta_2 = 0 \tag{B.5}$$

as well. Hence the two noise models are asymptotically equivalent. Note that for $|\overline{s}_{k,i}-\overline{s}_{k,i}| \gg \sigma$, the size of $b-a$ required to achieve convergence is exceedingly large, due to the rapid decay of the Gaussian function and the nature of loglikelihood. This is not an issue, however, since in this

case the statistical separability of the targets must be very high for the random deletions model, thus creating a low probability of error. Hence it is not important that the expected loglikelihoods converge so long as the $b - a$ is large enough that the separability is sufficiently large in order to maintain an absolute accuracy in predicted probability. This criterion is much easier to achieve, thus making the approximation more useful.

Bibliography

[1] U. Grenander, A. Srivastava, and M. Miller, "Asymptotic performance analysis of bayesian target recognition," *Information Theory, IEEE Transactions on*, vol. 46, no. 4, pp. 1658 –1665, jul 2000.

[2] B. Yen and J. Shapiro, "Asymptotic target recognition performance for flir and ladar systems," *Proc. of SPIE*, vol. 4379, no. 1, 2001.

[3] J. H. Dixon and A. D. Lanterman, "Information-theoretic bounds on target recongnition performance from laser radar data," *Proceedings of SPIE*, vol. 6234, no. 623419, 2006.

[4] D. R. Gerwe and P. S. Idell, "Cramrrao analysis of orientation estimation: viewing geometry influences on the information conveyed by target features," *J. Opt. Soc. Am. A*, vol. 20, no. 5, pp. 797–816, 2003.

[5] U. Grenander, M. Miller, and A. Srivastava, "Hilbert-schmidt lower bounds for estimators on matrix lie groups for atr," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 8, pp. 790 –802, aug 1998.

[6] M. Lindenbaum, "Bounds on shape recognition performance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 7, pp. 666 –680, jul 1995.

[7] A. Koksal, J. Shapiro, and M. Miller, "Performance analysis for ground-based target orientation estimation: Flir/ladar sensor fusion," in *Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on*, vol. 2, 1999, pp. 1240 –1244 vol.2.

[8] A. Jain, P. Moulin, M. Miller, and K. Ramchandran, "Information-theoretic bounds on target recognition performance based on degraded image data," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 9, pp. 1153 – 1166, sep 2002.

[9] M. H. DeGroot, *Probability and Statistics*, 2nd ed. Addison Wesley Publishing Company, January 1986.

[10] H. S. Kim and A. Hero, "When is a maximal invariant hypothesis test better than the glrt?" in *Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on*, vol. 1, 2000, pp. 401 –405 vol.1.

[11] Q. Zheng, S. Der, and H. Mahmoud, "Model-based target recognition in pulsed ladar imagery," *IEEE Transactions on Image Processing*, vol. 10, no. 4, April 2001.

[12] S. DelMarco, E. Sobel, and J. Douglas, "Hierarchical searching in model-based ladar atr using statistical separability tests," in *Automatic Target Recognition XVI*, F. Sadjadi, Ed., vol. 6234, no. 623403, 2006.

[13] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego: Academic Press, 1990.

[14] F. Garber and E. Zelnio, "On some simple estimates of atr performance, and initial comparison for a small data set," in *Proceedings of SPIE*, vol. 3070, 1997.

[15] M. Vore, "Analytical performance evaluation of sar atr with inaccurate or estimated models," in *Algorithms for Synthetic Aperture Radar Imagery XI*, E. Zelnio and F. Garber, Eds., vol. 5427, (2004).

[16] A. Jelalian, *Laser Radar Systems*. USA: Artech House, 1992.

[17] D. Jacob, P. Gatt, and T. Nichols, "Overview of lmct's advanced ladar simulator (alass)," *Proceedings of SPIE*, vol. 6950, no. 69500L, 2008.

[18] T. Chevalier and O. Steinvall, "Laser radar modeling for simulation and performance evaluation," in *Electro-Optical Remote Sensing, Photonic Technologies, and Applications III*,

G. Kamerman, O. Steinvall, K. Lewis, R. Hollins, T. Merlet, G. Bishop, and J. Gonglewski, Eds., vol. 7482, no. 748206, 2009.

[19] D. Youmans, "Avalance photodiode detection statistics for direct detection laser radar," in *Laser Radar VII*, 1992.

[20] M. Oh, H. Kong, and T. Kim, "Systematic experiments for proof of poisson statistics on direct-detection laser radar using a geiger mode avalanche photodiode," *Current Applied Physics*, vol. 10, 2010.

[21] S. Der, B. Redman, and R. Chellapa, "Simulation of error in optical range measurements," *Applied Optics*, vol. 36, no. 27, September 1997.

[22] P. Cho, H. Anderson, R. Hatch, and P. Ramaswami, "Real-time 3d ladar imaging," in *35th Applied Imagery and Pattern Recognition Workshop*, 2006.

[23] P. Gatt and S. Henderson, "Laser radar detection statistics: A comparison of coherent and direct detection receivers," in *Laser Radar Technology and Applications VI*, G. W. Kamerman, Ed., vol. 4377, 2001.

[24] B. Aull, J. Burns, C. Chen, B. Felton, H. Hanson, C. Keast, J. Knecht, A. Loomis, M. Renzi, A. Soares, V. Suntharalingam, K. Warner, D. Wolfson, D. Yost, and D. Young, "Laser radar imager based on 3d integration of geiger-mode avalanche photodiodes with two soi timing circuit layers," in *IEEE International Solid-State Circuits Conference*, 2006.

[25] R. Marino, W. Davis, G. Rich, J. McLaughlin, E. Lee, B. Stanley, J. Burnside, G. Rowe, R. Hatch, T. Square, L. Skelly, M. O'Brien, A. Vasile, and R. Heirichs, "High-resolution 3d imaging laser radar flight test experiments," *Proc. of SPIE*, vol. 5791, no. 138, 2005.

[26] S. Jordan, E. Armstrong, H. Larsson, M. Gebhardt, and O. Steinvall, "Scene-based algorithm for range/intensity estimation correction for the flash 3d ladar system," *Proc. of SPIE*, vol. 7684, no. 76840Q, 2010.

[27] J. Green and J. T. Shapiro, "Maximum-likelihood laser radar range profiling with the expectation-maximization algorithm," *Optical Engineering*, vol. 31, no. 11, pp. 2343–2354, 1992.

[28] J. Green and T. J. Shapiro, "Maximum-likelihood laser radar range profiling with the expectation-maximization algorithm," *Optical Engineering*, vol. 33, no. 3, pp. 865–874, 1994.

[29] G. Tolt and et al., "Registration and change detection techniques using 3d laser radar data from natural environments," in *Proceedings of SPIE*, vol. 6396, no. 63960A, (2006).

[30] L. Sjoqvist, C. Gronwall, M. Henriksson, P. Jonsson, and O. Steinvall, "Atmospheric turbulence effects in single-photon counting time-of-flight range profiling," *Proc. of SPIE*, vol. 7115, no. 71150G-1, 2008.

[31] T. J. Green, "Target recognition theory for laser radar imaging," *Proc. SPIE*, vol. 3070, p. 200, 1997.

[32] J. H. Dixon, "Ladar simulator user manual, version 0.7," Unpublished, January 2007.

[33] L. Tierny and J. B. Kadane, "Accurate approximations for posterior moments and marginal densities," *J. Amer. Statist. Assoc.*, vol. 81, pp. 82–86, 1986.

[34] B. Javidi, A. Fazlollahi, P. Willet, and P. Refregier, "Performance of an optimum receiver designed for pattern recognition with nonoverlapping target and scene noise," *Applied Optics*, vol. 34, no. 20, pp. 3858–3868, July 1995.

[35] N. Mehta, J. Wu, A. Molisch, and J. Zhang, "Approximating a sum of random variables with a lognormal," *Wireless Communications, IEEE Transactions on*, vol. 6, no. 7, pp. 2690 –2699, july 2007.

[36] L. Fenton, "The sum of log-normal probability distributions in scatter transmission systems," *IRE Transactions on Communications Systems*, 1960.

[37] A. Lanterman, M. Miller, and D. Snyder, "General metropolis-hastings jump difffusions for automatic target recognition in infrared scenes," *Opt. Eng.*, vol. 36, no. 4, pp. 1123,1137, April 1997.

[38] C. Gronwall, T. Chevalier, G. Tolt, and P. Andersson, “An approach to target detection in forested scenes,” in *Laser Radar Technology and Application XIII, Proc. SPIE*, vol. 6950, no. 69500S, (2008).