ALGORITHMS AND PERFORMANCE OPTIMIZATION FOR DISTRIBUTED RADAR AUTOMATIC TARGET RECOGNITION

A Dissertation
Presented to
The Academic Faculty

by

John S. Wilcher

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2015

ALGORITHMS AND PERFORMANCE OPTIMIZATION FOR DISTRIBUTED RADAR AUTOMATIC TARGET RECOGNITION

Approved by:

Dr. Aaron Lanterman, Advisor
School of ECE
Georgia Institute of Technology

Dr. William L. Melvin, Advisor
Georgia Institute of Technology
Georgia Tech Research Institute

Dr. Mark A. Richards
School of ECE
Georgia Institute of Technology

Dr. Jennifer E. Michaels
School of ECE
Georgia Institute of Technology

Dr. Douglas B. Williams
School of ECE
Georgia Institute of Technology

Dr. David A. Bader
School of CSE, College of Computing
Georgia Institute of Technology

Date Approved: March 4, 2015

For my wife, Diana;

our children, Brittany, Blaise, Katarina, and Stephanie;

and, my parents, Bettye and Jeff.

ACKNOWLEDGMENTS

The work detailed in this thesis would not have been accomplished without the care, support, understanding, and passion of many individuals. Without question, my deepest appreciation goes to my family – wife Diana; and children, Brittany, Blaise, Katarina, and Stephanie. The sacrifices each of them has made so I could pursue my passion are numerous. Some say it takes a village to raise a child; well, it takes a strong and loving family to support a mid-life PhD student. I will be forever indebted to each of them for helping me reach my goal.

Of course, the seeds of my passion for engineering were planted long ago by my mother and father. Throughout my life, they have encouraged me to perform at the highest level, regardless of my endeavor. Their support during my early college years was fundamental in ensuring I remained on a path to success. Their excitement and interest during my years at the Navy provided the spark for what has turned out to be a successful engineering career. Thanks for letting me take stuff apart and put it back together. I suspect it was necessary to get where I am today. Thank you, Mom and Dad.

I want to extend special appreciation to my friend Audrey Paulus. As is often the case, it is very helpful to have a partner when tackling a difficult task. Your support and advice throughout this process has been extremely helpful.

This journey would not have been possible without the help provided by Dr. Mark Richards. Dr. Richards met with me even before I started a single class at Georgia Tech. We spent several sessions discussing my past, the rigors of a PhD at Georgia Tech, and my love for military engineering. His class on radar signal processing was the best class I could have taken to start my journey.

Finally, I need to thank my advisors Dr. Bill Melvin and Dr. Aaron Lanterman for the countless hours they've spent mentoring me. They respected the challenges presented by a fully-employed PhD student and always sought to provide encouragement when the "chips were down." Their relentless pursuit for excellence in my work has helped me prepare for the demands associated with a career in research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This thesis focuses upon automatic target recognition (ATR) with radar sensors. Recent advancements in ATR have included the processing of target signatures from multiple, spatially-diverse perspectives. The advantage of multiple perspectives in target classification results from the angular sensitivity of reflected radar transmissions. By viewing the target at different angles, the classifier has a better opportunity to distinguish between target classes. This dissertation extends recent advances in multi-perspective target classification by: 1) leveraging bistatic target reflectivity signatures observed from multiple, spatially-diverse radar sensors; and, 2) employing a statistical distance measure to identify radar sensor locations yielding improved classification rates.

The algorithms provided in this thesis use high range resolution (HRR) profiles – formed by each participating radar sensor – as input to a multi-sensor classification algorithm derived using the fundamentals of statistical signal processing. Improvements to target classification rates are demonstrated for multiple configurations of transmitter, receiver, and target locations. These improvements are shown to emanate from the multi-static characteristics of a target class' range profile and not merely from non-coherent gain. The significance of dominant scatterer reflections is revealed in both classification performance and the "statistical distance" between target classes. Numerous simulations have been performed to interrogate the robustness of the derived classifier. Errors in target pose angle and the inclusion of camouflage, concealment, and deception (CCD) effects are considered in assessing the validity of the classifier. Different transmitter and receiver combinations and low signal-to-noise ratios are analyzed in the context of deterministic, Gaussian, and uniform target pose uncertainty models. Performance metrics demonstrate

increases in classification rates of up to 30% for multiple-transmit, multiple-receive platform configurations when compared to multi-sensor monostatic configurations.

A distance measure between probable target classes is derived using information theoretic techniques pioneered by Kullback and Leibler. The derived measure is shown to suggest radar sensor placements yielding better target classification rates. The predicted placements consider two-platform and three-platform configurations in a single-transmit, multiple-receive environment. Significant improvements in classification rates are observed when compared to ad-hoc sensor placement.

1 INTRODUCTION

While automatic target recognition can be performed with other sensors, this thesis focuses upon ATR with radar. The abundance of radar data and the complexities in using this data for target classification prompted the development of ATR techniques as early as the 1960s. ATR algorithms accept radar sensor data and attempt to match this data to known target characteristics stored within a reference library. The success of an ATR algorithm hinges on its ability to discern between distinct features of probable target classes. The distinctiveness of a target's features depends upon many factors including the target's physical characteristics, the operational environment, the sensitivity of the radar platform's sensors, and the signal processing capabilities of the radar platform. ATR is a difficult problem for the radar community; hence, significant effort has been expended for decades to advance ATR technology.

1.1 ATR Challenges

ATR technology must overcome a number of technological, environmental, and operational obstacles to be effective. The most challenging of these obstacles can be summarized as follows:

1. Any minor variation in scene geometry (e.g., platform or target movement) produces a significant change in the electromagnetic (EM) signature received at a radar sensor;

2. Due to scene-dependent variability in target characteristics and the number of potential target classes, libraries of known target characteristic data may be large. These large data sets induce computer storage, processing power, and processing time constraints; and,

3. The required target recognition rates must be extremely high to prevent misclassification errors that could result in significant, unintended property damage and personal injury.

These challenges necessarily impact the design of the radar system and, more specifically, the classifier. Variation in scene geometry significantly impacts the classifier's ability to discern between probable target classes. The computing requirements dictated by the large target characteristic database result in an ever-present engineering design trade-off between radar platform size and signal processing capacity. Finally, the need for such high target classification accuracy limits satisfactory operational conditions to those consistent with usable signal-to-noise ratio (SNR).

1.2 HRR-based Classifiers

Since ATR's inception, the radar community has experimented with a variety of target identification concepts. In recent decades, the majority of these efforts have focused upon the use of either HRR profiles or synthetic aperture radar (SAR) images as inputs to target classification algorithms. Successful ATR algorithms have been achieved with both sensor input types. SAR inputs have the advantage of high resolution, but are hindered by image blur for moving targets [1]. HRR inputs have the disadvantage of providing only a one-dimensional perspective of the target, while enjoying reduced computational and data collection requirements [2]. In this study, HRR profiles are the input of choice due to the desire to reduce the platforms' signal processing and data collection requirements.

High range resolution profile-based ATR algorithms rely on some form of pattern recognition. With these algorithms, the HRR profiles are submitted to the classification routine in

one of two forms. The profiles are either submitted without modification or subjected to various methods of feature extraction. Some of the feature extraction methods used include principle component-based approaches [1, 3, 4], rough set theory [5], central moments [6], distribution entropy [6], and non-orthogonal feature vectors [7]. Since feature extraction approaches require pre-processing prior to classifier input, potentially involving the loss of information, this thesis uses unmodified HRR profiles as input to the classifier to reduce the signal processing load on the radar platforms.

Regardless of the input format, algorithms employed by HRR profile-based classifiers rely on statistical or structural techniques to perform pattern matching [8]. Some of the most common statistical and structural techniques include Bayesian [9, 10], correlation [11], maximum likelihood [12], support vector machines [13, 6], nearest neighbor [6], and feed-forward artificial neural network [14, 15].

Another major distinction among HRR-based classifiers involves the introduction of multiple radar platforms. When inputs from multiple radar platforms are available for target identification, the classifier algorithm can accept all inputs for processing (sensor-level fusion) or perform classification on a sensor-by-sensor basis prior to final decision processing (decision-level fusion). Many tradeoffs are associated with deciding between one-stage or two-stage classification. For example, decision-level fusion has the advantage of reduced communication bandwidth between participating radar platforms and the disadvantage of missing potential data correlations among the different platforms' target views. Examples of successful decision-level fusion classifiers include those developed by Vespe [14, 15], Rizvi [16], and Kahler [17]. In contrast, sensor-level fusion has the potential to leverage multi-view data correlations with the disadvantage of having to transmit sensor data to a central processing platform. Successful

sensor-level fusion classifiers include those developed by Zhang [18], Nguyen [10], and Shaw [1]. A single-stage classifier employing sensor-level fusion was selected for this study to reduce the processing requirements of the platforms operating near the target of interest and avoid "throwing away" potentially useful data. This implies there is a more capable central processing platform in the vicinity of the less computationally-capable platforms. Power not used for local computation may be needed to transmit larger data sets. This thesis focuses on the advantages provided by more principled data analysis, leaving detailed study of the power tradeoffs for future work.

An analysis of previous works related to target classification from multiple target views suggests the inclusion of additional target perspectives helps the classifier identify a target's probable class. Specifically, studies performed by Vespe [14, 15, 19], Cui [20], and Zhang [18] demonstrate classifier sensitivity to scene geometry can be exploited by classifier algorithms to better discern between potential target classes. In addition, research by Stinco [11], and Butler [21] has shown the influence of multi-static reflectivity on classifier performance. These studies suggest additional target classification performance gains are possible if multiple, spatially-diverse platforms leverage multi-static reflectivity. The remainder of this thesis presents evidence of performance gains achievable when multiple, distributed radar platforms consider multi-static reflectivity in target classification.

1.3 Key Contributions

The key contributions of this research effort are:

1) The development of a multi-static, distributed, multi-sensor classifier algorithm that demonstrates better classification performance when compared to existing single and multi-perspective classifiers; and,

2) The development of a technique for identifying radar sensor placement that demonstrates improved classification performance.

Contributions provided by our classifier algorithm advance the state-of-the-art research of Shaw [1], Vespe [15], Kahler [17], Cui [20], Nguyen [10], Wong [9], and Zhang [18] by considering multi-static reflectivity. The combination of spatial diversity and sensor-level fusion represents advancement over the sensor-level fusion techniques provided by Nguyen [10], Shaw [1], Butler [21], and Kahler [17]. Based upon an extensive search of published research, little work appears to have been performed on the relationship between multiple sensor locations and classification. The research detailed in this thesis on the relationship between sensor location and improved classification significantly expands the work performed by Vespe [14] and Zhang [18].

This thesis provides a framework for addressing multi-static, distributed, multi-sensor ATR. This framework leverages the accomplishments of Jacobs [12] on using sequences of HRR profiles for target classification. In [12], the authors demonstrate the validity of using HRR profiles to perform ATR based upon a conditional Gaussian data model. This model serves as the basis for all work presented in this thesis. The conditional Gaussian data model states that target HRR profiles exhibit independent Gaussian behavior conditioned on a specific target aspect angle and target class. The legitimacy of the conditional independence is rooted in the understanding that target returns are uncorrelated when separated by sufficient angular extent. This model is used in conjunction with the findings of Vespe [15] on multi-perspective classification. Vespe demonstrated progressive improvement in classification rates as additional target perspectives were added to the classifier. These results were proven using three different classifiers, thereby suggesting the ATR improvements were generally due to additional perspectives and not a specific classifier algorithm. By extending the work performed by Jacobs and Vespe to multi-

static target returns, we demonstrate how the uniqueness of a target's multi-static reflectivity can be leveraged for additional gains in ATR performance.

Once the benefits of multi-static reflectivity are demonstrated, a detailed analysis of radar sensor location and classification is presented. Information-theoretic techniques are employed to derive a lower bound on a divergence measure attributable to target class. This divergence measure lower bound is used to identify sensor locations where target classes are most discernible. Based entirely upon the classifier derived in the first part of this research effort, the divergence measure bound is compared to probabilities of correct classification over a range of noise levels to demonstrate a correspondence between target class divergence and classification performance. Algorithms incorporating the divergence measure bound are presented to demonstrate how the bound can be used to identify platform positions in single transmitter, multiple receiver classification scenarios.

Throughout this thesis, the terms sensor and platform are used interchangeably. Either instance refers to an airborne vehicle capable of transmitting and/or receiving EM waves suitable for further radar processing.

2 BACKGROUND

Modern radar systems serve many purposes within government, commercial, and military communities. The design, development, deployment, and operation of these systems vary based upon their intended use within each community. However, despite their various forms, all radar systems perform one or more of the following functions: detection, tracking, classification, and imaging. These functions leverage EM wave transmission and reflection characteristics as key discriminants in their respective signal processing algorithms. More specifically, radar systems process reflected EM energy to: 1) recognize the presence of desired objects (detection); 2) follow the movement of a detected object (tracking); 3) identify the object as belonging to a particular object type (classification); and, 4) produce a one, two- or three-dimensional representation of the object (imaging). We focus our treatment of reflected EM wave energy to classification processing. Radar detection, tracking, and imaging functions are assumed to exist and support the classification function as required.

A block diagram of a typical radar system (adapted from [22]) showing the major radar receive processing components is provided in Figure 1. EM reflections from the radar's operating environment are captured by the antenna and amplified by the radio frequency amplifier (RFA). Once amplified, the received signal is modulated to a lower frequency by the mixer. Additional amplification is provided by the intermediate frequency amplifier (IFA) before the signal is delivered to the radar's signal processor. The signal processor is responsible for several key functions prior to classification processing. These key functions may include: 1) sampling, 2) matched filtering, and 3) detection processing. The design of the radar's antenna, front-end electronics, analog-to-digital converters, filters, and detector are critical to radar performance, including classification. Design decisions related to these radar components dictate

many key radar performance measures including signal-to-interference plus noise ratio (SINR), range resolution, probabilities of detection and false alarm, and percentages of correct target classification.

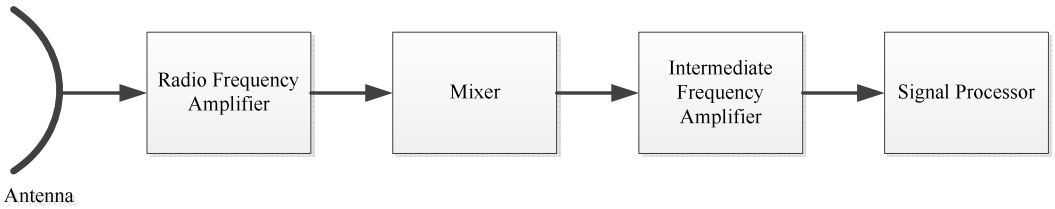


Figure 1. Typical radar receive processing block diagram.

Our classification study assumes the pre-classification signal processing components have been designed to support the required classifier data input requirements. The division between "signal processor" and "classifier" is somewhat arbitrary; we make it to emphasize the classification aspects.

2.1 Problem Formulation

Radar systems transmit EM waveforms and process waveform energy reflected from objects in the radar's operating environment. Reflected energy can emanate from any physical object present within the transmitted waveform's beam width. We restrict our research to reflections from ground-based vehicles.

In its most general form, our standard signal processing model, not including noise, entails the convolution of the transmitted waveform with the range-scattering function of the target:

$$y(t) = \int_{-\infty}^{\infty} \gamma x(t - t_\lambda) b(\lambda)\, d\lambda, \tag{2.1}$$

where $x(t)$ is the transmitted waveform; $b(\lambda)$ is the range-scattering function; $y(t)$ is the waveform returned to the radar platform; γ represents the cumulative losses incurred by the waveform in the environment; t is time; and, t_λ is the time associated with distance λ. The range-scattering function depends on distance, but not time. This singular dependence on distance has resulted in the description of illuminated targets as range-spread targets [23]. A representative diagram of (2.1) is shown in Figure 1 (adapted from [24]).

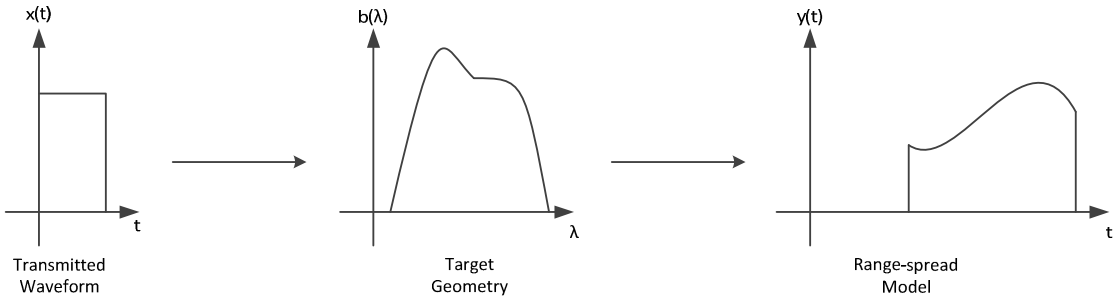


Figure 2. Diagrammatic representation of transmitted waveform convolution with a target.

The transmitted waveform impinges on numerous surfaces of the target and reflects the transmitted waveform in directions dictated by the azimuthal and elevated normals of the surfaces. This process causes some waveform energy to be returned to the radar platform, while other energy is directed away. The reflected waveforms received by the radar platform are dependent upon the target's physical properties.

This thesis focuses on the classification of ground-based targets. Since all targets are illuminated with the same transmitted waveform and are assumed to experience the same

environmental losses, target uniqueness is completely determined by the range-scattering function. Identification of unique scattering characteristics of the observed targets is required to properly classify the targets. The scattering characteristics of targets are collectively referred to as radar cross section (RCS). As will be shown in Section 2.4, a target's RCS is highly dependent upon the angular position of the target relative to the radar platform (target pose angle) and the incident and reflected angles of the transmitted waveform. Coupled with a target's unique physical properties, the reflected waveform geometrical dependencies provide an opportunity to distinguish between two targets that might otherwise look very similar. This serves as the premise for leveraging multiple, dispersed radar platforms to more reliably identify a target's class.

2.2 Waveforms

The waveforms transmitted by radar systems are designed to enhance the detection, tracking, classification, and imaging of a desired scene or objects. The specific design characteristics of these waveforms are formulated based upon the intended use of the radar system. For this study, a waveform suitable for providing excellent range resolution and simplicity in post-receipt processing is desirable. A popular choice for meeting these requirements is a pulsed waveform with frequency modulation. The pulsed waveform permits our radar system to transmit and receive from a single platform while the frequency modulation characteristic ensures the system provides sufficient range resolution. A pulsed waveform with frequency modulation can be expressed in complex notation as:

$$x(t) = rect(t) \cdot A(t) \exp\left\{ j\left(2\pi f_c t + \theta(t)\right)\right\}, \tag{2.2}$$

where $rect(t) = 1 \text{ for } 0 \le t \le \tau$ and $rect(t) = 0 \text{ for } t < 0 \text{ and } t > \tau$; $A(t)$ is the amplitude modulation; exp is the exponential function; j is $\sqrt{-1}$; f_c is the carrier frequency; $\theta(t)$ is the phase function that includes the requisite frequency modulation; and, τ is the pulse length. The complex exponential form is algebraically convenient; the "real" waveform is found by taking the real part of $x(t)$. It is rare for the transmitted waveform to include amplitude modulation, so $A(t)$ is assumed to be constant throughout the pulse's duration. Typical pulsed radar systems employ several pulses to achieve their objectives. A multi-pulse version of (2.2) is:

$$\begin{aligned} x_{pulsed}(t) &= \sum_{m=0}^{M-1} x(t - mT) \\ &= \sum_{m=0}^{M-1} rect(t - mT) \cdot A \exp\left\{ j\left(2\pi f_c (t - mT) + \theta(t - mT)\right)\right\} \end{aligned}, \quad (2.3)$$

where m is the pulse number, M is the total number of pulses, and T is the pulse repetition interval (PRI). This waveform has been designed to use M pulses within a single coherent processing interval (CPI) of length $M \times PRI$.

Several forms of frequency modulation are possible with the waveform described in (2.3). Linear, non-linear, stepped, and linear-stepped frequency are the most prevalent frequency modulation forms used by today's radar systems. Linear and linear-stepped frequency modulation waveforms are often used to generate HRR profiles similar to those required by this study. In certain operational scenarios, a stepped frequency waveform can also be used to generate the HRR profiles. The results of Chapters 4 and 5 are applicable to all HRR profiles regardless of the waveform type used.

2.3 High Range Resolution Profiles

The HRR profiles used in this study are simulated using a simple complex-valued reflectivity waveform. This approximation reduces the complexity and processing time required to generate the required HRR profiles by simulating the output of a typical frequency modulated waveform signal processing chain. As detailed in Section 2.2, several different frequency modulated waveforms are suitable for HRR profile generation. The profiles can be generated using matched filtering, which results in the output of a radar processor appearing like the HRR convolved with the autocorrelation of the transmitted waveform. Some waveforms, such as stepped-frequency waveforms, facilitate computationally efficient approaches using the fast Fourier transform. Linear FM waveforms permit other tricks such as "stretch processing," which can reduce receiver bandwidth requirements. Our approach does not require any particular waveform or technique; we merely assume that the effective point-spread response of the resulting HRR formation process can be reasonably modeled as a Dirac delta function, and that such functions are summed into "range bins" aligned with a resolution of

$$\Delta R = \frac{c}{2\beta} \ , \tag{2.4}$$

where β is the frequency bandwidth of the transmitted waveform and c is the speed of light (approximated as 3×10^8 meters per second).

It has been shown [25] that the scatterer-reflected waveform received at a radar platform, within a proportionality constant, is

$$\overline{y}(t) = \sum_{h=1}^{H} \sqrt{\left|\sigma_h\left(\theta_t, \theta_r, \theta_{pose}\right)\right|} \exp\left\{ j2\pi f \left(t - \frac{2R_h\left(\theta_t, \theta_r, \theta_{pose}\right)}{c} \right) \right\} \ , \tag{2.5}$$

where $\sqrt{\left|\sigma_h\left(\theta_t,\theta_r,\theta_{pose}\right)\right|}$ represents the square root of the RCS magnitude for scatterer h; $R_h\left(\theta_t,\theta_r,\theta_{pose}\right)$ is the distance between the radar platform and scatterer h; f is the center frequency of the waveform; θ_t is the angle of the transmit waveform relative to the scatterer; θ_r is the angle of reflection relative to the scatterer; θ_{pose} is the target pose angle relative to the same scene geometry imposed on the scatterer; and, H is the number of scatterers. With the appropriate substitution of distance for time in (2.5), the discrete, scatterer-reflected waveform received at the radar platform depends upon the position and RCS of the scatterers

$$y[n]=\sum_{h=1}^{H_n}\sqrt{\left|\sigma_h\left(\theta_t,\theta_r,\theta_{pose}\right)\right|}\exp\left\{2\pi f\left(\frac{2R_{n,h}\left(\theta_t,\theta_r,\theta_{pose}\right)}{c}\right)\right\}, \tag{2.6}$$

where $R_{n,h}$ is the range to scatterer h that is present in range bin n and H_n is the number of scatterers in range bin n. If observed for a specific slow-time sample, the scatterer-reflected waveform becomes

$$\begin{aligned} y[n,m]=\sum_{h=1}^{H_n}&\sqrt{\left|\sigma_h\left(\theta_t[m],\theta_r[m],\theta_{pose}[m]\right)\right|} \\ &\cdot\exp\left\{2\pi f\left(\frac{2R_{n,h}\left(\theta_t[m],\theta_r[m],\theta_{pose}[m]\right)}{c}\right)\right\}, \end{aligned} \tag{2.7}$$

where $\theta_t[m],\theta_r[m]$, and $\theta_{pose}[m]$ are now dependent upon the position of the platform during each slow-time sample m. The slow-time sample m is not meant to imply that coherent processing is going to take place across the m dimension.

2.4 Radar Cross Section

Radar cross section is as a measure of the reflective strength of a target defined as 4π times the ratio of the power per unit solid angle scattered in a specified direction to the power per unit area in a plane wave incident on the scatterer from a specified direction [26] [27]. In mathematical form, the definition can be expressed as

$$\sigma = \lim_{r \to \infty} 4\pi R^2 \frac{\left| \vec{E}^{scattered} \right|^2}{\left| \vec{E}^{incident} \right|^2} , \tag{2.8}$$

where σ designates RCS measured in square-meters (sm), R is the distance between the radar transmitter and the target, $\vec{E}^{scattered}$ is the electric field scattered from the target measured in volts, and $\vec{E}^{incident}$ is the electric field impinging upon the target measured in volts. The limit with respect to R addresses the need to standardize RCS measurements regardless of distances between the radar transmitter, target, and receiver. Intuitively, RCS can be described as the product of three components [27]

$$\sigma = \text{Projected Cross Section} \times \text{Reflectivity} \times \text{Directivity}, \tag{2.9}$$

where

Projected Cross Section	Geometrical area subjected to the illuminating waveform,
Reflectivity	Portion of the transmitted power reradiated by the target, and
Directivity	Ratio of the power scattered towards the radar receiver relative to power scattered isotropically.

As demonstrated in (2.9), RCS is not solely a function of the physical cross section exposed to the radar transmitter. Target construction, including material properties and geometrical attributes, plays a significant role via the reflectivity and directivity components.

An understanding of RCS is critical to this study as demonstrated in (2.7). Section 2.1 argued that the ability to classify targets is strictly dependent upon each target's range-scattering function. The range-scattering function presented in (2.7) was simplified for purposes of clarity. A more complete formulation of target RCS must consider all of the factors listed below:

- Transmit angle - θ_t,
- Receive angle - θ_r,
- Target orientation relative to transmitter/receiver - θ_{pose},
- Target construction (materials and geometrical composition),
- Transmit waveform frequency - f, and
- Transmitter and receiver polarization - $S^{transmit}$, $S^{receive}$.

Several of these effects are not necessary to demonstrate the technology outlined in this thesis. Specifically, the effects of polarization are not addressed. Thus, the use of a scattering matrix and other polarization concepts are excluded from the analysis; this could be an avenue for future work. Additionally, no consideration for the target's material properties has been included. However, target geometrical construction and transmit waveform frequency must be considered due to the target model requirements outlined in Section 3.1. To accommodate the dependencies on target geometrical construction and waveform frequency, (2.7) is expanded in a more complete form as

$$\sqrt{\left|\sigma_h\left(\theta_t[m],\theta_r[m],\theta_{pose}[m],f,\mathbf{d}(h)\right)\right|}, \tag{2.10}$$

where $\mathbf{d}(h)$ is a vector of parameters describing dimensional aspects of the scatterer. This parameter set may vary depending upon scatterer type (sphere, point, plate, etc.). For our test targets, we use point and plate scatterers to form representative target classes. However, the technology presented in Chapter 4 and Chapter 5 could be used on target reflectivity data generated from EM codes, including material properties, for example.

An important consideration in the specification of target RCS is the scattering regime applicable to the target. There are three regimes used to describe EM scattering. The regimes are defined in terms of target size and EM wavelength. Targets with physical dimensions much smaller than the transmitted wavelength fall within the Rayleigh region. If the transmitted wavelength is roughly the size of the target, scattering occurs within the resonant region. Finally, if the transmitted wavelength is much smaller than the target, the scattering is defined to be within the optics region. The region designation is important in determining which of the numerous EM scattering types are applicable. Examples of the various scattering types include traveling, specular, end-region, diffraction, and multiple bounce. Given the transmit wavelength and target characteristics used in this study, RCS scattering will be restricted to the optics region. Operating within the optics region requires careful consideration of the specular, end-region, diffraction, and multiple bounce scattering types.

Multiple theories have been developed to model scattering from objects in the high frequency (optics) region. Examples include geometric optics, physical optics, geometrical theory of diffraction (GTD), uniform theory of diffraction (UTD), method of equivalent currents (MEC), physical theory of diffraction (PTD), and other extensions based on these models [28]. Each of these models is subject to some restriction in geometric modeling due to incident angle, reflective angle, or both. This study employs the GTD models developed in [29] to generate the

required RCS values. These GTD models support the parametric requirements outlined in this section. Further details on the use of these models to generate RCS are covered in Section 3.1. Again, it is important to emphasize the technology presented in Chapter 4 and Chapter 5 is not dependent upon any specific method for modeling RCS.

2.5 Detection and Classification Theory

Detection theory encompasses the concepts and methods necessary to determine the presence of a signal in noise. Detection theory becomes classification theory when the detection algorithms seek to identify a particular signal in noise as being a particular signal type. For example, speech classification algorithms attempt to recognize specific sounds or words from an audio recording. In the context of target recognition, classification algorithms endeavor to characterize received target signals as belonging to a particular target class. These problems are categorized as multiple hypothesis testing problems. Typically, these testing problems involve comparisons of observed data to available, known data sets exhibiting probable target class features.

The complexity of the classification problem is directly related to what is known about the probability density functions (PDF) of all potential target classes and the observed HRR profiles. Since the signal representations defined later in Section 3.2 are inherently random, statistical hypothesis testing must be employed. Also, since the statistical models defined in Chapter 3 are parameterized by unknown attributes (e.g. target pose angle), a composite hypothesis testing approach must be used to derive the necessary classifier. Most statistical, composite hypothesis testing approaches result in classifiers obtained from one of two general strategies: generalized likelihood ratio tests (GLRT) and Bayesian tests. The GLRT has the binary testing form [30]

$$L_G(\mathbf{z}) = \frac{p\left(\mathbf{z};\hat{\boldsymbol{\theta}}_i, H_i\right)}{p\left(\mathbf{z};\hat{\boldsymbol{\theta}}_j, H_j\right)} > 1, \tag{2.11}$$

where $\mathbf{z}$ represents the data vector under test; $\hat{\boldsymbol{\theta}}_i$ and $\hat{\boldsymbol{\theta}}_j$ are the maximum-likelihood estimates (MLE) for the unknown parameters that maximize $p\left(\mathbf{z};\hat{\boldsymbol{\theta}}_i, H_i\right)$ and $p\left(\mathbf{z};\hat{\boldsymbol{\theta}}_j, H_j\right)$, respectively; and, the detection threshold present in [30] has been replaced by 1 to signify a classification versus detection problem. The probable target class is determined to be i if (2.11) is satisfied.

The Bayesian approach has the binary testing form [30]

$$\frac{p(H_i \mid \mathbf{z})}{p(H_j \mid \mathbf{z})} = \frac{p(\mathbf{z} \mid H_i)\, p(H_i)}{p(\mathbf{z} \mid H_j)\, p(H_j)} = \frac{p(H_i) \int p(\mathbf{z} \mid \boldsymbol{\theta}_i, H_i)\, p(\boldsymbol{\theta}_i \mid H_i)\, d\boldsymbol{\theta}_i}{p(H_j) \int p(\mathbf{z} \mid \boldsymbol{\theta}_j, H_j)\, p(\boldsymbol{\theta}_j \mid H_j)\, d\boldsymbol{\theta}_j} > 1, \tag{2.12}$$

where $p(\mathbf{z} \mid H_i)$ is the PDF of $\mathbf{z}$ with respect to hypothesis H_i; $p(\mathbf{z} \mid \boldsymbol{\theta}_i, H_i)$ is the conditional PDF of $\mathbf{z}$ conditioned upon $\boldsymbol{\theta}_i$ for hypothesis H_i; $p(\boldsymbol{\theta}_i; H_i)$ is the prior PDF for the parameter $\boldsymbol{\theta}_i$ under hypothesis H_i; and, $p(H_i)$ is the probability of target class H_i. Assuming each target class is equally likely and the prior PDFs on pose are the same regardless of target class, the Bayesian binary test form reduces to

$$\frac{p(\mathbf{z} \mid H_i)}{p(\mathbf{z} \mid H_j)} = \frac{\int p(\mathbf{z} \mid \boldsymbol{\theta}_i, H_i)\, p(\boldsymbol{\theta}_i)\, d\boldsymbol{\theta}_i}{\int p(\mathbf{z} \mid \boldsymbol{\theta}_j, H_j)\, p(\boldsymbol{\theta}_j)\, d\boldsymbol{\theta}_j} > 1. \tag{2.13}$$

Assuming the classes are equally likely, the probable target class is determined to be i if (2.12) is satisfied. The classifier derived in Section 4.1 uses the Bayesian approach.

The multiple hypothesis testing approach required by this study uses (2.12) to decide between C probable target classes. The probable target class is determined by evaluating

$$\hat{c} = \arg\max_{c} p\left(z \mid H_c\right), \tag{2.14}$$

where $\hat{c}$ is the probable target class; $c, \hat{c} \in C$; and, C is the index set of all potential target classes.

2.6 Performance Metrics

2.6.1 Probability of Correct Classification

The primary performance metric for this study is probability of correct classification (PCC). PCC denotes the percentage of test HRR profile samples identified as the correct target class

$$PCC_i = \frac{\sum_{s} I\left(p\left(\mathbf{z}_s \mid H_i\right) > p\left(\mathbf{z}_s \mid H_j\right)\right)}{S_{total}}, \tag{2.15}$$

where PCC_i is the percentage of correct classification for target class i; $I(\cdot)$ is the indicator function; $\mathbf{z}_s$ is the HRR profile associated with sample index s; and, S_{total} is the total number of HRR profiles submitted to the classifier. The results of the classification tests are stored in confusion matrices to document how well the classifier performs for all possible target classes. A typical confusion matrix is provided in Table 1.

Table 1. Typical confusion matrix showing performance percentages of test data against a library of known target data for three probable target classes.

		Training Target Class		
		1	**2**	**3**
Test Target Class	**1**	75	15	10
	2	20	80	0
	3	15	15	70

Two PCC metrics are derived from the confusion matrix and used in performance plots. First, a mean PCC value is calculated by taking the average of the confusion matrix diagonal. This metric represents the overall performance of the classifier referenced to properly classifying all possible target classes. An example of this metric is the comparison of monostatic and multi-static classifier performance as noted in Figure 15. A second metric analyzes the PCC values of a single confusion matrix row. This second metric references classifier performance specific to a single test target class. An example of the second metric is classification results related to CCD use as noted in Figure 25.

2.6.2 SNR

Throughout this thesis, performance plots are provided where the horizontal axis denotes noise power. Noise power was chosen to ensure consistency when comparing results across multiple tests and plots. Due to Rayleigh fading, target construction, and other test configuration characteristics, consistency in SNR could not be assured for all transmit incidence angles, receive platform angles, and target classes. Noise power increments of 3 dB were employed in the generation of reflectivity data according to (3.3). In almost all cases, the respective single-pixel

SNR values range from approximately 28 to 33 dB for a noise power of 0 dB. The referenced SNR values were calculated using an average of the zero-noise-power target reflectivity amplitude across slow-time samples and within a range window that only included the target. This average amplitude was converted to decibels to give the approximate SNR levels mentioned above.

2.7 Approach

The air-to-ground scenario addressed by this study simulates an operational environment where a single, stationary, non-cooperative target is interrogated by several spatially-distributed radar platforms. Each platform illuminates the target in a straight-line flight path and collects HRR profile data accordingly. The HRR profile data is submitted to a centralized classifier for classification processing wherein a forced decision model determines the probable target class. This data collection and processing architecture is depicted in Figure 3.

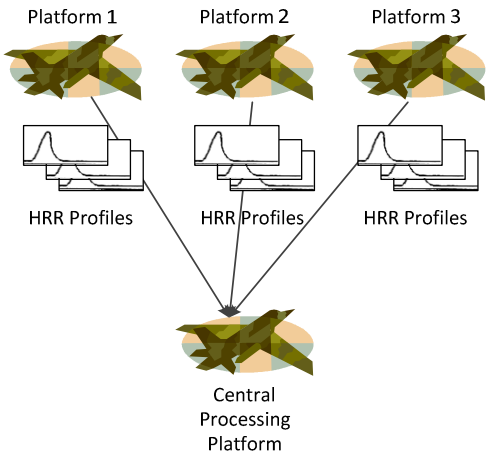


Figure 3. Distributed data collection and centralized data processing architecture.

The scenario described above has been formulated to address several key elements of this study:

1. The selection of HRR profiles for sensor data supports a general goal of reducing the computational requirements on the distributed platforms;

2. Multiple sensors provide the flexibility to interrogate the target of interest from multiple perspectives at the same time; and,

3. A centralized classification processing platform reduces the need for significant processing power on the distributed platforms and avoids the potential loss of useful sensor data.

A physics-based scattering model suitable for the generation of realistic, multi-static target return data is developed in Chapter 3. This scattering model is useful because of the limited public availability of multi-static target return data. Also, our scattering model allows us to intuitively modify scattering effects to better interpret how such modifications affect performance results. Chapter 4 derives our multi-sensor classifier using the conditional Gaussian data model. In contrast to several earlier multiple-perspective studies that selected classifiers of interest [13, 15, 16, 20], statistical decision theory dictates the form of our classifier. The classifier's final form leverages the independence of each transmit and receive sensor combination to obtain a composite test statistic. Numerous data sets are generated to compile a library of test and training data samples. These data samples are parameterized by the transmit incidence, receive platform, and target pose angles in addition to target class.

Once separated into test and training data libraries, the data sets are subjected to Monte Carlo simulations employing several platform and target configurations. The platform test configurations are defined specifically to evaluate the advantages and disadvantages of using

multiple, distributed sensors at the same time. Specifically, the benefits and drawbacks associated with having one or more transmitting platforms operating in conjunction with one or more receiving platforms is demonstrated. For example, these tests lead to a performance assessment comparison between two intriguing platform configurations: 1) a single transmitting platform with three receiving platforms leveraging multi-static target returns; and, 2) three monostatic platforms operating independently. The case of all platforms transmitting and receiving is also of interest.

The impact of target pose uncertainty on classifier performance is examined for three target pose uncertainty models. These tests demonstrate the importance of selecting the appropriate target pose angle prior probability distribution relative to classifier operating conditions. Different test target configurations are explored in terms of concealment, camouflage, and deception characteristics. These tests evaluate the robustness of the classifier design to incidental and purposeful modification of known target classes. The results of these studies indicate the conditions where specific platform configurations perform better under CCD challenges. All test results are analyzed with particular attention to low SNR performance.

After validating the merits of the distributed, multi-sensor classifier, the impact of sensor placement relative to classification performance is analyzed in Chapter 5. Information-theoretic concepts are used to derive an angle-dependent distance measure. This distance measure considers scene geometry (transmit, receive, and target pose angles) to identify the configurations for which the target classes differ the most. We leverage an approximation of the Kullback-Leibler divergence (KLD) to determine scene geometry providing the greatest target class probability density separation. Two algorithms are developed to enumerate transmit and receive angles producing the greatest KLD values. Each algorithm assumes three platforms (one transmit

and two receive) are available for interrogating the target. The first algorithm presumes all three platforms operate independently and thus, finds the highest KLD value with three degrees of freedom. The second algorithm determines the highest KLD value for a two platform (one transmit and one receive) scenario. Using the selected transmit angle, a second receive angle is identified using the second highest KLD value. This second algorithm represents a significant reduction in processing when compared to the first algorithm. The KLD results are compared to the commensurate PCC results to validate algorithm performance.

Finally, an important terminology distinction may help avoid misunderstandings. Much work has been performed to understand the benefits of multiple-input, multiple-output (MIMO) radar systems [31, 32, 33]. Many of these MIMO studies have addressed the benefits of spatial diversity. These are, to the best of my knowledge, related to detection and possibly geolocation. MIMO radar systems typically involve cooperation among radar platforms in terms of signal transmission and receipt. The radar platforms employed in this study cooperate in the sense of sharing sensor data but are independent in transmission and receipt processing. A notable exception to platform independence is the understanding that each platform transmits a waveform that is orthogonal to, and therefore, easily distinguishable from the others in terms of target return processing.

3 DATA MODELS

Examining the impact of spatially-diverse target scattering on ATR performance requires a robust target reflectivity data set parameterized by transmitter incidence, receiver platform, and target pose angles. Publicly-available data sets exhibiting spatially-diverse target scattering are rare. Thus, a mechanism to produce synthetic target reflectivity data is helpful for analyzing the behavior of a multi-static, distributed, multi-sensor classifier. The target reflectivity data supporting this study is generated using target models composed of individual scatterers. The subsections below describe the scatterer-based target and signal modeling used to generate the required HRR profiles. The methods of Chapters 4 and 5 could also be applied to training data derived from EM codes such as XPATCH, Lucernhammer, compact radar range test, or carefully calibrated studies with real targets.

3.1 Target Model

Targets are simulated as a collection of non-dominant and dominant scatterers formed in the canonical shape of an ellipse. Scatterer density and location are parameterized in a manner to support the notion of an electrically-large target. Electrically-large targets have reflectivity characteristics associated with the optical scatterering regime. A more imposing constraint of the target design assumes the distances between two adjacent scatterers are, in almost all cases, larger than the transmit waveform's wavelength. This high-frequency constraint ensures the HRR profile received at any radar platform exhibits typical optical regime scattering properties such as specular scattering and diffraction [28, 34]. The additional optical regime scattering properties of multiple bounce and end-region scatterering are not considered in this study. The specular

scattering and diffraction properties were sufficient to illustrate the benefits of bistatic scattering on ATR performance.

Three target models are constructed to simulate targets from three distinct target classes. Every target class has the same scatterer positions but varying scatterer RCS values. Non-dominant scatterers have RCS values ranging from 10 to 15 dBsm within different regions of each target class. The exact scatterer RCS specifications are provided in Section 4.2.2. Dominant scatterers exhibit distinctive characteristics as discussed further below. Scatterer positions are determined by distributing individual scatterers within a 0.5 meter block according to a uniform distribution. The scatterer density is controlled to ensure the majority of scatterer-to-scatterer distances are greater than the transmit waveform's wavelength. The complete target is finalized by combining sixteen 0.5 meter blocks into the roughly elliptical shape referenced earlier. The final target form provides a two-dimensional representation suitable for generating plausible target reflectivity at reasonable target-to-platform distances. A third dimension (elevation) was not essential to demonstrating the aforementioned ATR performance improvements. The impact of incorporating elevation into target and signal modeling may be addressed in future work. A typical target representation is provided in Figure 4, where red circles indicate locations of non-dominant scatterers with individual RCS values of 10 dBsm. Similarly, the blue and green circles indicate non-dominant scatterer locations with RCS values of 13 and 15 dBsm, respectively. The solid black circles identify the locations of the four dominant scatterers.

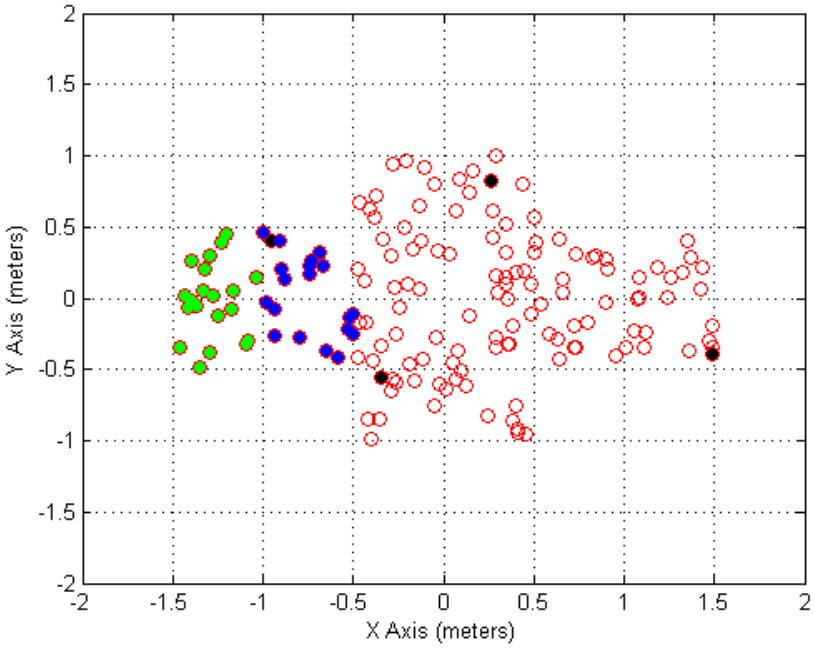


Figure 4. Typical target representation showing the density and locations of target scattering centers.

Non-dominant scatterers are modeled as spheroids. The RCS of the non-dominant scatterers model perfect reflection of incident EM energy. Four dominant scatterers are distributed in four relatively dispersed locations within the targets. The dominant scatterers are modeled as flat plate reflectors and placed within each target class to ensure distinctiveness between target classes as a function of pose angle. All plate reflections within a specific target class have the same normals. The plate normals are 18, 36, and 45 degrees relative to the negative y-axis for target classes 1, 2 and 3, respectively. Each plate is modeled as square with a length of 0.5 meters and a peak RCS of 30 dBsm. This work models flat plate radar cross section using GTD principles as provided in [35]:

$$\sigma_{dom}\left(f,\phi_t,\phi_r,\theta_t,\theta_r;L_m;h_m\right)= \left(j\frac{f}{f_c}\right)\mathrm{sinc}\left(\frac{\pi f}{c}L_m\left(\sin\theta_t\cos\phi_t+\sin\theta_r\cos\phi_r\right)\right)\times \mathrm{sinc}\left(\frac{\pi f}{c}h_m\left(\sin\phi_t+\sin\phi_r\right)\right), \quad (3.1)$$
$$\phi_t,\phi_r,\theta_t,\theta_r\in\left[-\pi/2,\pi/2\right]$$

where σ_{dom} is the dominant scatterer RCS; ϕ_t is the elevation transmit angle; ϕ_r is the elevation receive angle; θ_t is the azimuthal transmit angle; θ_r is the azimuthal receive angle; f is transmit waveform frequency; f_c is the transmit waveform carrier frequency; L_m is the length of the flat plate; and, h_m is the flat plate height. The angles ϕ_t and ϕ_r are assumed to be zero degrees, given our two-dimensional assumption. An example RCS plot for the flat plate is provided in Figure 5. Note the significant drop in RCS as the receive angle moves away from the maximum specular reflection angle. In this example, RCS drops by more than 20 dB with a deviation of only one degree from the plate normal.

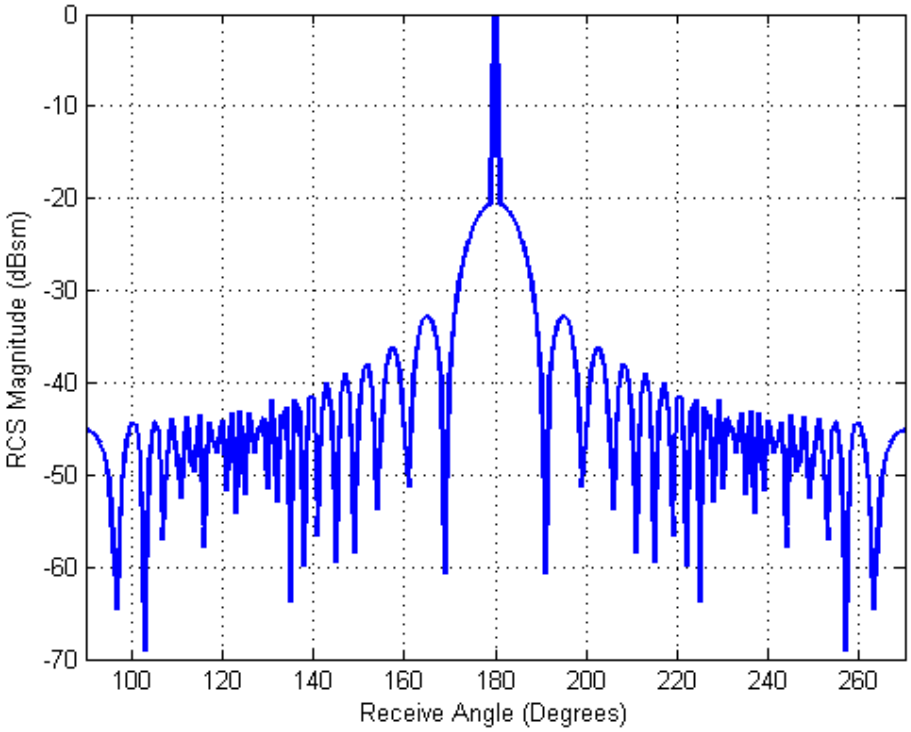


Figure 5. Example plot showing flat plate RCS response at a frequency of 10 GHz, transmit angle of 0 degrees, plate dimensions of 0.5 meters x 0.5 meters, and plate normal of 90 degrees.

3.2 Signal Model

Each target scatterer collection is used to determine the composite radar target return at a specific receive platform. The composite radar target return simulates the HRR profile that would be received from transmitting a wideband signal whose autocorrelation is approximated as an impulse, as described in Section 2.3. The transmit waveform is assumed to be configured to support the necessary energy and range resolution requirements. The composite radar target return signal received by an individual radar platform is

$$y[n,m] = \sum_{h=1}^{H_n} \sqrt{\left|\sigma_h\left(\theta_t[m], \theta_r[m], \theta_{pose}[m], f_m, \mathbf{d}(h)\right)\right|} \cdot \exp\left\{2\pi f_m \left(\frac{2R_{n,h}\left(\theta_t[m], \theta_r[m], \theta_{pose}[m]\right)}{c}\right)\right\}. \quad (3.2)$$

This signal is combined with receiver noise to produce

$$z_{a,b}[n,m] = y_{a,b}[n,m] + v_b[n,m], \quad (3.3)$$

where $z_{a,b}$ is the waveform received at platform b due to the transmission from platform a; and, v_b is additive, circular-symmetric, white Gaussian noise (AWGN) with zero mean and variance σ^2_{noise}.

The signal represented in (3.3) is used to produce HRR profiles from individual scatterers according to the position of the applicable transmitters and receivers. The generated HRR profiles exhibit the expected characteristics of Rayleigh fading, range stretch, and range compression. Rayleigh fading is demonstrated in the HRR sequence plot of Figure 6, which shows sample target HRR profiles for 45 sequential illuminations separated by one meter in cross-range. Note the varying reflectivity amplitudes predominantly in range bins 12 through 20. Also, note that the fading in range bins 11 and 21 are relatively slight. This reduced fading is due to the dominant scatterer contributions exhibited by the flat plate reflectors. The flat plate reflector RCS response does not drop significantly due to the small angle traversed during the data collection (approximately 0.25 degrees).

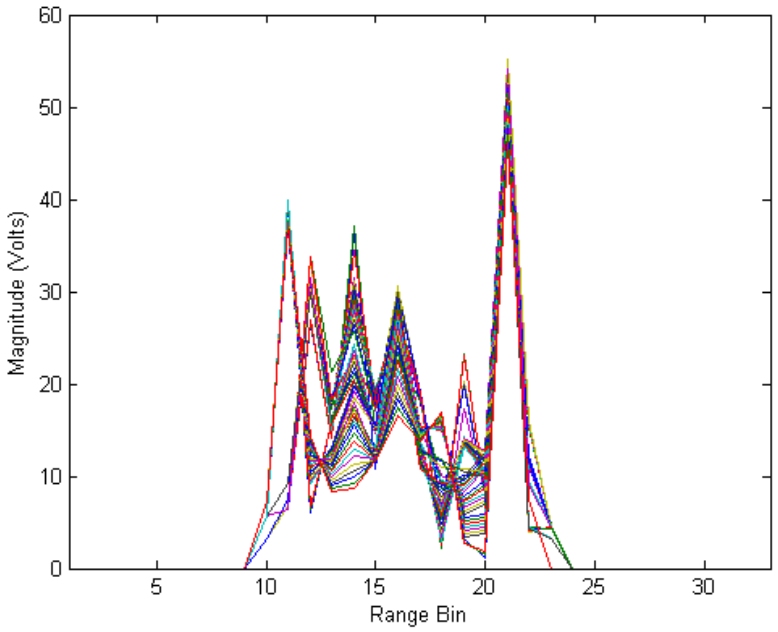
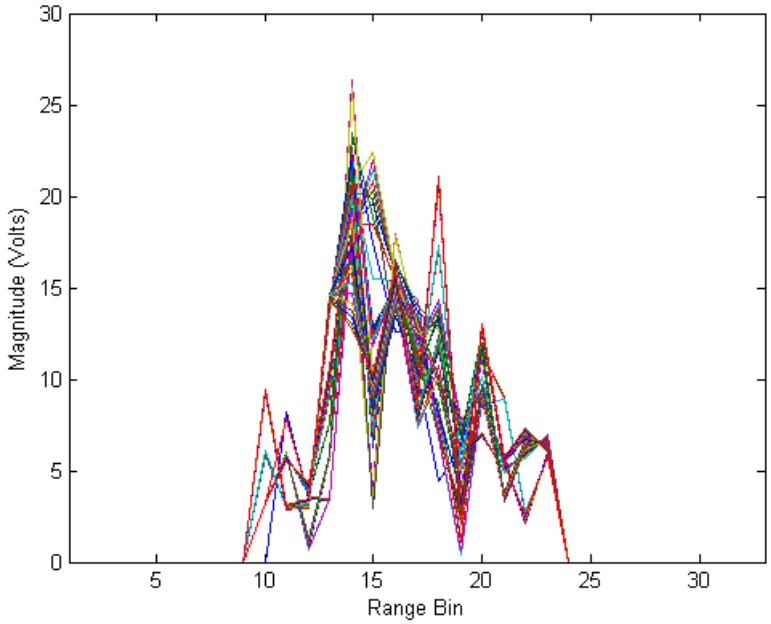
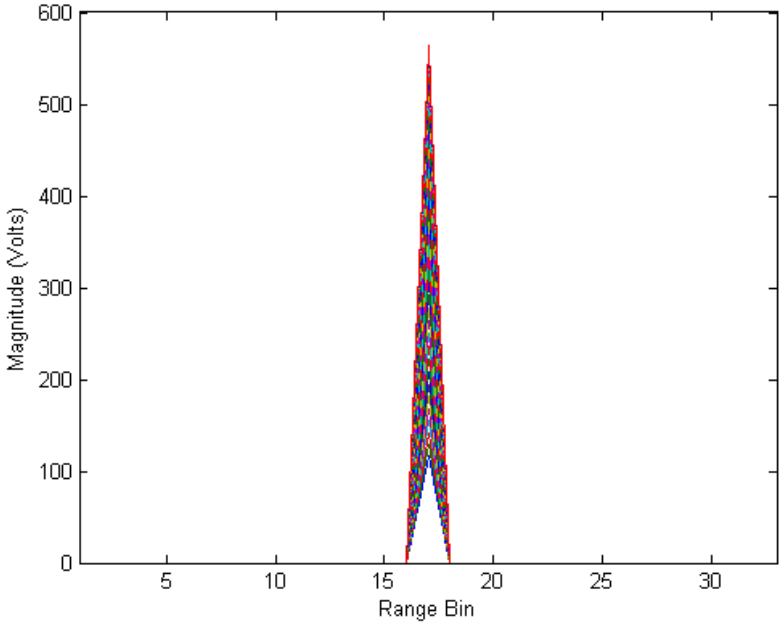


Figure 6. Sequence of HRR profiles showing Rayleigh fading.

Bistatic angle geometry induces a natural stretching and compression of the HRR profile based upon transmitter and receiver placement. HRR profiles are stretched when the bistatic angle is zero and severely compressed at bistatic angles approaching 180 degrees, as shown in Figure 7. The zero degree case is the monostatic or pseudo-monostatic case. The HRR profile for the bistatic angle of zero degrees is spread across 16 range bins, versus one range bin for the bistatic angle of 180 degrees. Also, note the significant increase in HRR profile amplitude when all scatterers are confined to the single range bin.



(a)



(b)

Figure 7. Comparison of HRR profile lengths at bistatic angles of 0 degrees (a) and 180 degrees (b).

Several simplifying assumptions have been applied to the data model, without loss of generality to the overall analysis. Specifically, occlusion of scatterers has not been modeled. This simplification allows scatterers behind illuminated flat plates to be included in the target's response. The inclusion of these scatterers in the target contributes more to the resulting constructive and destructive summation of individual scatterer responses. Multipath reflections are also not considered within the model. The complexities associated with modeling multipath have been excluded to reduce processing time. Due to the relatively large number of scatterers already included in the target response, any additional benefits of including multipath responses were deemed unnecessary for our research objectives. Additionally, the target class shape and individual scatterer positions are not intended to mimic any specific real world target. The target shape and scatterer positions were chosen to demonstrate the effects of spatial diversity on ATR.

Finally, a significant assumption about the statistical properties of the HRR profile is required to simplify classifier design in accordance with the research constraint of reducing processing complexity. Gaussian random processes afford many opportunities for simplifying the derivation of signal detectors and classifiers. These detectors and classifiers often result in more practical and efficient implementations as well. However, in the case of this classification study, limitations associated with a realistic training data set and scene geometry force us to consider a composite radar return whose magnitude may be Rician, as discussed below. Any phase included in the training data sets is problematic given the difficulty of ensuring that the test receiver is at the exact distance (or wavelength multiples) from the target as observed when the training data was collected. Because the exact phase of the radar return is difficult to use reliably within the

classifier, the magnitude of the composite radar return is used as input to the classifier. The magnitude operation, $\widehat{\mathbf{z}}_{a,b} = \left|\mathbf{z}_{a,b}\right|$, is performed on each discrete range bin as follows:

$$\widehat{\mathbf{z}}_{a,b}[m] = \left[\left|z_{a,b}[1,m]\right| \;\; \left|z_{a,b}[2,m]\right| \;\; \ldots \left|z_{a,b}[N,m]\right|\right]^T. \tag{3.4}$$

Because the distribution of $\mathbf{z}_{a,b}$ is complex Gaussian with additive dominant scatterers, the distribution of $\widehat{\mathbf{z}}_{a,b}$ is Rician. We approximate $\widehat{\mathbf{z}}_{a,b}$ as Gaussian, and not Rician, to simplify classifier implementation in accordance with the practicality goal stated in Chapter 1.

In [12], the received complex data were assumed to be circularly-symmetric zero-mean Gaussian random variables, and conditionally independent between range bins. Under these assumptions, the magnitudes of data are Rayleigh distributed, and they have a convenient mathematical form. Although we sometimes make the same assumptions when interpreting the behavior of our algorithms, our procedures do not depend on these assumptions; we fit a real-valued Gaussian distribution with non-zero mean and non-diagonal covariance to the magnitudes of the training data. Adding noise before taking the magnitude allows us to map the underlying density on the complex data through the magnitude transformation analogous to the way random sampling is employed in particle filters or Monte Carlo integration in general. If the original complex data does have some correlations between range bins, the resulting density would have a multivariate Rayleigh distribution, which is analytically foreboding. If it also has non-zero mean with uniformly distributed phases, it would have a multivariate Rician distribution, which is even more complicated [36], and contains nonstandard integrals. Rather than try to compute these directly with techniques such as trapezoidal integration and then try to match Gaussian parameters to those densities, we find it reasonable and tractable to simply represent the density of the magnitudes as an ensemble of random samples and fit a Gaussian distribution to that

ensemble. The underlying range bin data may be uncorrelated, or not; it may be zero mean, or not; either way, our approach encapsulates those effects. This Gaussian approximation on the *magnitudes* of $\widehat{\mathbf{z}}_{a,b}$, which contrasts our "conditional Gaussian" implementation with the implementation pioneered in [12] that is applied to *complex values*, affects the results of this study only in the sense that the theoretical optimality of the derived classifiers is not assured. The results of Chapter 4 illustrate that the Gaussian approximation leads to reasonable results.

4 MULTI-SENSOR, DISTRIBUTED, MULTI-STATIC ATR CLASSIFIER

This chapter presents the theoretical formulation of the multi-sensor, distributed, multi-static classifier. Our classifier derivation employs an approach similar to the one used by Jacobs [12] to model HRR profiles as conditionally Gaussian as opposed to the more complex Rician form. The classifier is subjected to several test scenarios for purposes of examining its performance in the context of monostatic classification, dominant scatterer contributions, target pose angle uncertainty, and robustness with respect to unknown target modifications. Tests are also conducted to demonstrate the sensitivity of classification results on radar platform placement. The later tests serve as motivation for development of the sensor placement algorithms presented in Chapter 5.

4.1 Classifier Design

The data model employed by our ATR algorithm leverages the conditional Gaussian model detailed in [12], except we fit a real-valued Gaussian to the magnitude data instead of fitting a complex-valued Gaussian to the original complex data. Under our conditional Gaussian model assumption, the waveform received at each radar platform under hypothesis H_i exhibits Gaussian behavior with a mean and covariance parameterized by the transmit incidence and receive angle vector $\boldsymbol{\theta}_{a,b} = \left[\theta_a^{transmit} \;\; \theta_b^{receive} \right]^T$ and target pose angle θ_{pose},

$$\widehat{\mathbf{z}}_{a,b}\left(\boldsymbol{\theta}_{a,b}, \theta_{pose} \mid H_i\right) \sim N\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right), \mathbf{C}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\right)^{1}, \qquad (4.1)$$

[1] When used to the right of a tilde, N denotes a Gaussian distributed random variable. When used as part of a mathematical equation, N represents the Gaussian density.

where $i \in C$ designates a specific target class; $\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)$ and $\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)$ are the mean and variance, respectively; θ_{pose} is the target pose angle. An example configuration of transmit platform, receive platform, and target pose angles is provided in Figure 8.

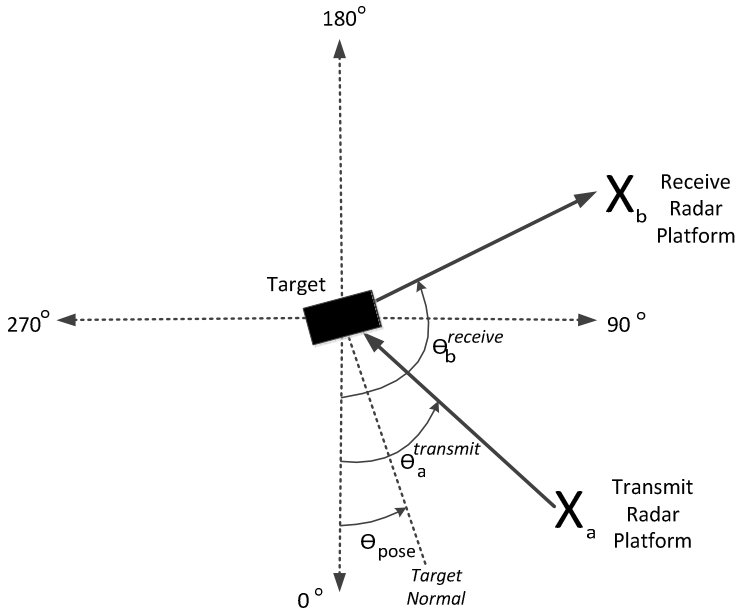


Figure 8. Depiction of radar platform and target pose angles.

Since $\widehat{\mathbf{z}}_{a,b}$ is parameterized by $\boldsymbol{\theta}_{a,b}$ and θ_{pose}, the signal model for $\widehat{\mathbf{z}}_{a,b}$ is described by a family of densities. As detailed in [37], statistical decision theory dictates a composite hypothesis testing approach to formulate the required detector/classifier. Without loss of generality, we presume the target classes are equally likely; hence, using the Bayesian likelihood ratio test (LRT), the classification decision between target class i and target class j has an initial form of

$$\frac{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid H_j;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}$$
$$= \frac{\int p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid \theta'_{pose}, H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)p\left(\theta'_{pose} \mid H_i\right)d\theta'_{pose}}{\int p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid \theta''_{pose}, H_j;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)p\left(\theta''_{pose} \mid H_j\right)d\theta''_{pose}} > 1 \quad , \tag{4.2}$$

where $p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)$ is the multi-platform, joint probability density of $\hat{\mathbf{z}}_{a,b}$ for hypothesis H_i given the specific transmit incidence and receive platform angles for each transmit and receive platform combination, $\boldsymbol{\theta}_{1,1}\cdots\boldsymbol{\theta}_{A,B}$; $p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid \theta_{pose}, H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B},\right)$ is the multi-platform, conditional probability density of $\hat{\mathbf{z}}_{a,b}$ for hypothesis H_i given the specific transmit incidence and receive platform angles for each transmit and receive platform combination, $\boldsymbol{\theta}_{1,1}\cdots\boldsymbol{\theta}_{A,B}$, and target pose angle, θ_{pose}; $p\left(\theta_{pose} \mid H_i\right)$ is the prior probability density of the target's pose angle θ_{pose} for hypothesis H_i; A is the number of transmit platforms; and, B is the number of receive platforms. Independent target pose angle integrations are required in (4.2) and separately notated as θ'_{pose} and θ''_{pose} to avoid confusion.

As stated in [12], each receive platform waveform, $\hat{\mathbf{z}}_{a,b}$, when conditioned on target type and scene geometry, can be considered statistically independent of other receive platform waveforms, provided the platforms are separated by sufficient angular extent. We assume the receive platforms are sufficiently spaced in angular extent to ensure the returns of two adjacent platforms are independent. This conditional independence permits the separation of the multi-platform, conditional probability density into the product of the individual densities as follows

$$p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid \theta_{pose}, H_i; \boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right) = p\left(\hat{\mathbf{z}}_{1,1} \mid \theta_{pose}, H_i; \boldsymbol{\theta}_{1,1}\right) \cdots p\left(\hat{\mathbf{z}}_{A,B} \mid \theta_{pose}, H_i; \boldsymbol{\theta}_{A,B}\right). \quad (4.3)$$

The density $p\left(\theta_{pose}; H_i\right)$ describes prior knowledge of the target's pose angle for each hypothesis. Without loss of generality, we assume each hypothesis has the same prior density on target pose angle. Our study considers three possible densities for a prior on the target pose: 1) impulse, 2) Gaussian,[2] and 3) uniform. The impulse density indicates exact knowledge of the target pose angle. The Gaussian density models the behavior of a typical tracking system, where the target pose angle is presumed known with some level of uncertainty. The uniform density assumes no prior knowledge of the target's pose angle. The final form of the LRT test statistic will differ based upon which prior density is considered. The test statistic derivation for the Gaussian prior is included below, followed by a simple statement of the remaining two test statistics.

Under the assumption of a Gaussian prior and conditional independence of the platform waveforms, the log LRT (LLRT) becomes

$$\ln\left\{\frac{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B} \mid \theta'_{pose}, H_i; \boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{AB} \mid \theta''_{pose}, H_j; \boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}\right\} = \ln\left\{\frac{\int \exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1} \mid \theta'_{pose}, H_i; \boldsymbol{\theta}_{1,1},\right)\cdots p\left(\hat{\mathbf{z}}_{A,B} \mid \theta'_{pose}, H_i; \boldsymbol{\theta}_{A,B}\right) p\left(\theta'_{pose}\right)\right]\right\} d\theta'_{pose}}{\int \exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1} \mid \theta''_{pose}, H_j; \boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B} \mid \theta''_{pose}, H_j; \boldsymbol{\theta}_{A,B}\right) p\left(\theta''_{pose}\right)\right]\right\} d\theta''_{pose}}\right\} > 0 \quad (4.4)$$

where $p\left(\hat{\mathbf{z}}_{a,b} \mid \theta_{pose}, H_i; \boldsymbol{\theta}_{a,b}\right) = N\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right), \mathbf{C}_i\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\right)$, $p\left(\theta_{pose}\right) = N\left(\theta_{mean}, \sigma_t^2\right)$;

θ_{mean} is the target pose angle estimated by some tracking algorithm; and, σ_t^2 is the tracking error

[2] Technically speaking, it is a "wrapped Gaussian" because we apply the Gaussian form to an angular parameter.

variance. Substituting the multivariate Gaussian density formulation into (4.4), replacing the continuous integration operation with the discrete sum,[3] removing constant terms appearing in each hypothesis test statistic, and simplifying the results gives a test statistic for each hypothesis as follows

$$T_k = \ln\left\{\sum_{\theta_{pose}}\left\{\exp\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left[\det\left(\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right)\right]\right.\right.\right.$$
$$\left.-\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right)^T\mathbf{C}_k^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right),\right. \tag{4.5}$$
$$\left.\left.\left.-\frac{\left(\theta_{pose}-\theta_{mean}\right)^2}{2\sigma_t^2}\right\}\right\}\right\}$$

where $k = i$ or j; and, $\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)$ and $\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)$ are the model mean and covariance for target class k at the prescribed transmit incidence, receive platform, and target pose angles.

Typically, one of the terms within the angle summation dominates, so we apply the approximation of using the maximum value across all angles, yielding the test statistic

$$T_k = \ln\left\{\max_{\theta_{pose}}\left\{\exp\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left[\det\left(\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right)\right]\right.\right.\right.\right.$$
$$\left.-\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right)^T\mathbf{C}_k^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}\right)\right)\right. . \tag{4.6}$$
$$\left.\left.\left.\left.\left.-\frac{\left(\theta_{pose}-\theta_{mean}\right)^2}{2\sigma_t^2}\right\}\right\}\right\}\right\}\right\}$$

The maximization in (4.6) can be changed to a minimization if the inequality in (4.4) and the signs in (4.6) are switched. In addition, since the log function is monotonically increasing we

[3] Without loss of generality, we assume that each target hypothesis is computing with the same set of angular samples approximating the integral, and that these samples are equally spaced.

can switch the order of the minimization operation and the log function to produce the final form of our LLRT:

$$T_i < T_j, \tag{4.7}$$

where

$$\begin{aligned} T_k = \min_{\theta_{pose}} \Bigg\{ \sum_{a=1}^{A} \sum_{b=1}^{B} \Big\{ & \ln\Big[\det\Big(\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\Big)\Big] \\ & + \left(\hat{\mathbf{z}}_{a,b} - \boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\right)^T \mathbf{C}_k^{-1}\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\left(\hat{\mathbf{z}}_{a,b} - \boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)\right), \\ & + \frac{\left(\theta_{pose} - \theta_{mean}\right)^2}{2\sigma_t^2} \Big\} \Bigg\} \end{aligned} \tag{4.8}$$

and the Gaussian superscript has been added to distinguish between each of the three pose angle prior density test statistics. Thus, the classifier selects a probable target class of i instead of j if the inequality in (4.7) is satisfied. Examining the form of (4.8), we note the presence of the Mahalanobis distance measure indicated by the quadratic term in $\hat{\mathbf{z}}_{a,b}$. The Mahalanobis distance provides a quantitative measure of how closely $\hat{\mathbf{z}}_{a,b}$ is related to the statistical probability density described by $\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)$ and $\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)$. This measure is shifted by the log of the determinant of $\mathbf{C}_k\left(\boldsymbol{\theta}_{a,b}, \theta_{pose}\right)$ and the target pose angle uncertainty. Intuitively, the covariance scale factor balances differences in power amongst the target classes. As can be seen in (4.8), the pose angle uncertainty scale factor increases the test statistic in instances where the pose angle deviates from the actual target pose angle.

The test statistics for the known (impulse density) and completely unknown (uniform density) pose angles follow easily from (4.8). If the exact pose angle is known, the test statistic reduces to

$$T_k^{Known} = \sum_{a=1}^{A}\sum_{b=1}^{B}\Big\{\ln\Big[\det\Big(\mathbf{C}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{actual}\Big)\Big)\Big] + \Big(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{actual}\Big)\Big)^T\mathbf{C}_k^{-1}\Big(\boldsymbol{\theta}_{a,b},\theta_{actual}\Big)\Big(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{actual}\Big)\Big)\Big\}, \tag{4.9}$$

where θ_{actual} is the actual target pose angle.

The test statistic associated with the unknown target pose angle prior density is

$$T_k^{Uniform} = \min_{\theta_{pose}}\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}\Big\{\ln\Big[\det\Big(\mathbf{C}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{pose}\Big)\Big)\Big] + \Big(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{pose}\Big)\Big)^T\mathbf{C}_k^{-1}\Big(\boldsymbol{\theta}_{a,b},\theta_{pose}\Big)\Big(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\Big(\boldsymbol{\theta}_{a,b},\theta_{pose}\Big)\Big)\Big\}\right\}. \tag{4.10}$$

4.2 Simulation

4.2.1 Simulation Environment

Consider an operational scenario where three radar platforms are available to sense a single, stationary ground-based target. Each platform is capable of transmitting and receiving waveforms whose transmission characteristics permit separation of returns from multiple transmitters (e.g. the transmission waveforms are orthogonal in some sense). The target is one of three distinct target classes, $c \in \{C\}$. The radar platforms and target are restricted to two dimensions, $\boldsymbol{\theta} = [\theta_t \;\; \theta_r]$, as noted in Section 3.1. An example of the scene geometry is provided in Figure 9.

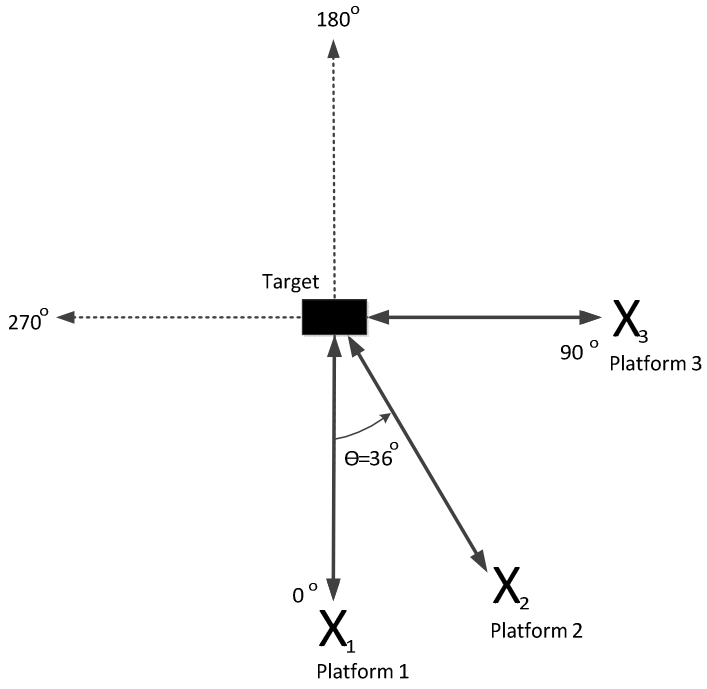


Figure 9. Scene geometry showing example placements of the target and three radar platforms.

The radar platforms operate in an environment without clutter and all propagation and other channel losses are assumed to be identical for all platforms and test scenarios. The illumination of the target by the radar platform occurs over a straight flight line path constituting a slow-time aperture distance of 90 meters. Slow-time samples[4] are captured every one meter. The distance between the center point of the slow-time aperture and the target is always 10,000 meters. The range and slow-time aperture dimensions have been selected to ensure all slow-time samples from the CPI lie within the correlation window as shown in Figure 10. The correlation angle extent is defined to be [38]:

[4] The term "slow-time" merely indicates a sequence of collected range profiles, and is not intended to indicate any kind of coherent processing over the slow-time dimension, such as in synthetic aperture radar or space-time adaptive processing.

$$\Delta\theta = \frac{c_{SoL}}{2Lf} = \frac{\lambda}{2L} \qquad (4.11)$$

where L is the maximum target length, c_{SoL} is the speed of light, f is frequency of the transmitted radar pulse, and $\lambda = c_{SoL} / f$ is the wavelength of the transmitted radar pulse. Illumination within the correlation angle extent is crucial for accurate statistical representation of the radar return distributions used in Section 4.1.

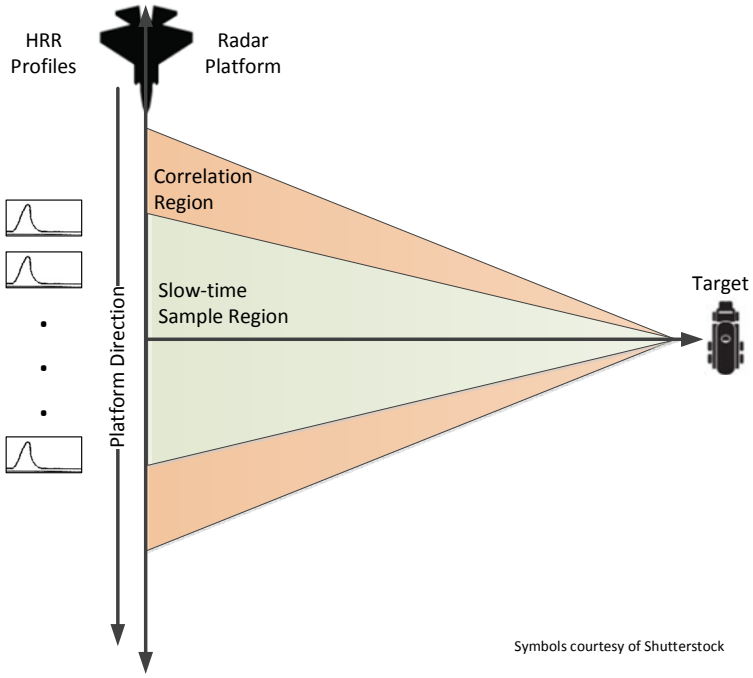


Figure 10. Slow-time aperture depiction showing presence of correlation and sample regions.

The complex-valued HRR profile obtained for each slow-time sample spans 10 meters divided into 33 range bins of 0.3 meters in length. The resulting data set from a single CPI is N x M, where $N = 33$ and $M = 90$. Test and training data samples are extracted from the data set

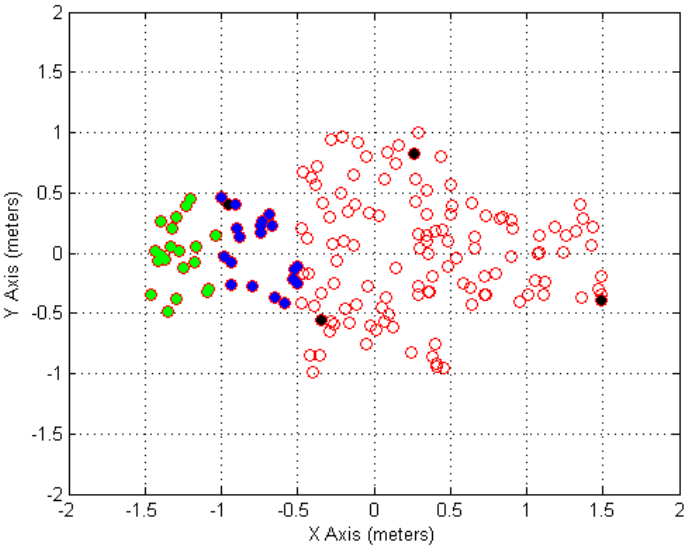
using an even-odd paradigm, producing an equal quantity of test and training range profiles. These independent data sets are separated and used accordingly in the classifiers defined in Section 4.1.

Each profile represented within a data set is the result of a bistatic transmission between a transmitter, target, and a receiver. The bistatic nature of the HRR profile necessitates the generation of HRR profiles for the full range of possible transmitter and receiver angles. Data sets were generated for 120 transmit and 120 receive angles, covering the full angular extent in 3-degree increments. In summary, 43,200 data sets of size $N \times M$ were generated to represent three target classes, 120 transmit angles, and 120 receive angles.

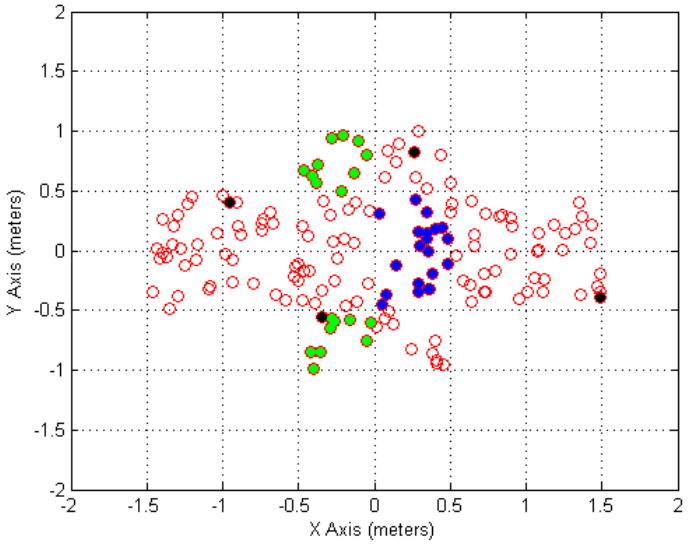
4.2.2 Target Construction

4.2.2.1 Non-dominant Scatterers

As indicated in Section 4.2, the target under interrogation represents one of three target classes. Each physical representation of a target class conforms to the scatterer model provided in Section 3.1. The non-dominant scatterer RCS values were varied across target class to ensure some level of distinctiveness. The non-dominant scatterer RCS distributions are shown in Figure 12. The red, blue, and green circles indicate RCS values of 10, 13, and 15 dBsm, respectively.

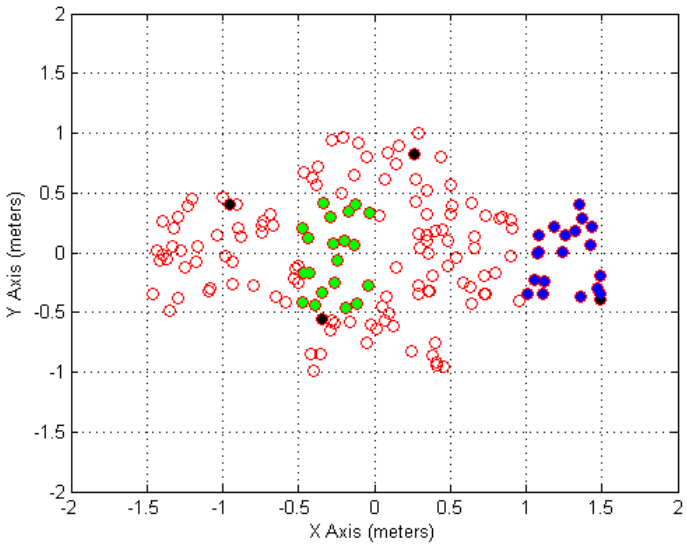


(a)



(b)

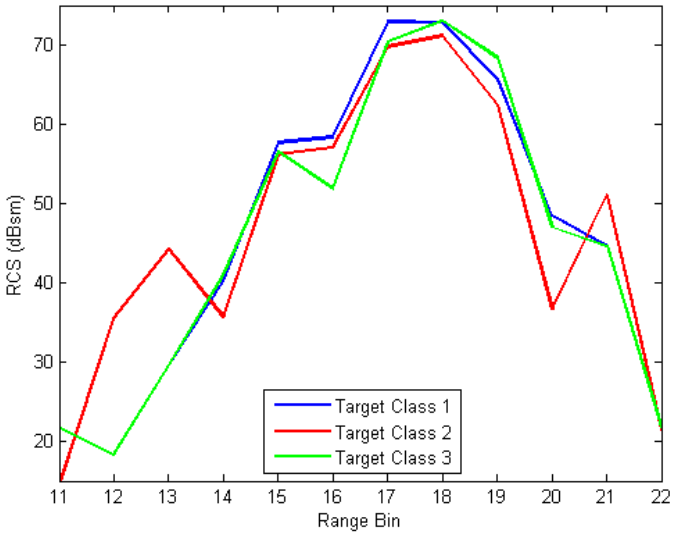
Figure 11. Target scatterer diagrams showing varying regions of RCS for target class 1 (a), target class 2 (b), and target class 3 (c).



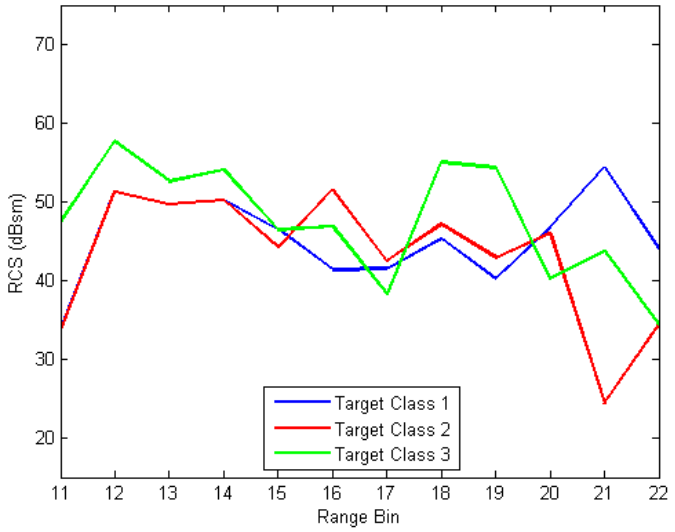
(c)

Figure 12. Continued.

The placement of the non-dominant scatterers with RCS values of 13 and 15 dBsm provide HRR profiles that demonstrate differences across target classes at some angles and similarity at other angles. An example of the HRR profile similarities and differences for specific combinations of transmit incidence and receive angles is shown in Figure 13. The plots have been windowed in range to remove range bins containing only noise.



(a)



(b)

Figure 13. Example windowed HRR profiles demonstrating a) similarities at a transmit angle of 0 degrees and a receive angle of 6 degrees; and, b) differences at a transmit angle of 0 degrees and a receive angle of 90 degrees.

4.2.2.2 Dominant Scatterers

The dominant scatterers within each target class play a significant role in demonstrating the benefits of multi-static, distributed, multi-sensor ATR. Dominant scatterer positions are identified by the black circles in Figure 12 and have an RCS of 30 dBsm. The use of flat plates as dominant scatterers provides a straightforward illustration of dominant scatterer contributions to the HRR profiles. The placement and orientation of the flat plates for each target class is provided in Table 2. The plates within each target class are oriented identically, but have different normals when compared to the other target classes. These plate orientations ensure dominant scattering distinctiveness across target classes.

Table 2. Dominant scatterer plate normals by target class

Target Class	Number of Plates	Plate Normals (Degrees)
1	4	18
2	4	36
3	4	45

The plate normals of 18, 36, and 45 degrees have been selected to complement the sensor placements used in several of the test scenarios. The complimentary selection of plate normals and sensor placement supports requirements for analyzing the benefits of a multiple transmit and receive operational scenario. Specifically, as shown in Figure 14, scattering from the flat plates contribute to the HRR profiles at all three platforms. In the case of platforms 1 and 2, flat plate scattering from multiple target classes is achieved. The target classes most readily distinguishable by dominant scattering at each receive platform are listed in Table 3.

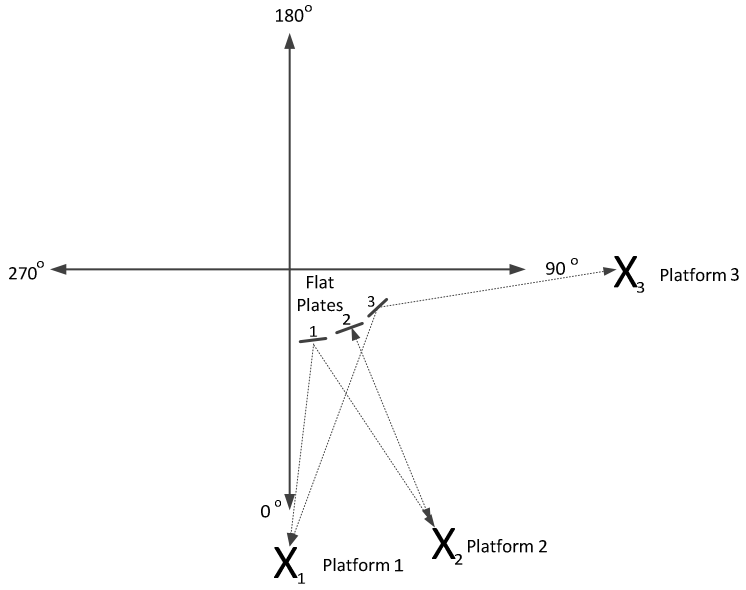


Figure 14. Multiple transmit and receive scene geometry.

Table 3. Target class identification by transmit and receive platform.

	Transmit Platform & Position		
Receive Platform	**1 (0 degrees)**	**2 (36 degrees)**	**3 (90 degrees)**
1	None	Target Class 1	Target Class 3
2	Target Class 1	Target Class 2	None
3	Target Class 3	None	None

4.2.3 Test Scenarios

The tests were designed to progressively demonstrate how multi-static reflectivity can be leveraged by multiple platforms. The tests validate the following key characteristics of the classifier:

1. The use of multi-static reflectivity with multiple platforms to improve classification performance is the result of multi-static reflectivity and not merely due to non-coherent gain achieved from using multiple platforms;

2. The presence of dominant scatterers in the multi-static returns increases classifier performance as more dominant scatterers are added;
3. Multi-static platform configurations outperform monostatic platform configurations;
4. The placement of the radar platforms impacts classifier performance;
5. The choice of probability models used for target pose uncertainty have some effect on classifier performance, but comes short of nullifying the benefits of the classifier, unless target pose angle is assumed as known; and,
6. The impact of the tested concealment, camouflage, and deception effects on the classifier is relatively minor.

The explored platform configurations are described in Table 4. Tests associated with test scenarios 1 and 3 are conducted with a fixed target pose angle of 0 degrees. All platform positions discussed in these sections are measured relative to this fixed target pose angle. The results presented in Sections 4.3.1 and 4.3.3 can be obtained for any fixed target pose angle, and any results related to platform positions would have their positions adjusted by the change in target pose angle.

Table 4. Test platform configurations and definitions.

Platform Configuration Number	Description	Purpose
1	Single Sensor, Monostatic	Provides a reference performance level simulating a single platform operational scenario
2	Multi-sensor, Monostatic	Demonstrates ATR performance improvements gained by using multiple radar platforms configured to process target returns associated only with each platform's transmission
3	Single Transmit, Multiple-receive, Bistatic	Analyzes ATR performance under the scenario of a single transmit platform (not co-located with the receive platforms) and multiple receive platforms
4	Multiple-transmit, Multiple-receive, Bistatic	Validates the ATR performance gains achieved when multiple platforms transmit and receive

4.2.3.1 Test Scenario 1

The simulations executed in this first test scenario are detailed in Table 5. These simulations use multiple platform configurations to establish several baseline classifier performance measures. Three radar platforms are initially placed such that each platform's transmit and receive angle is coincident with a specific target class dominant scatterer plate normal (test platform configurations 1 and 2). Radar platform one is positioned at 18 degrees to receive maximum response from the flat plate scatterers of target class 1. Similarly, radar

platforms two and three are placed at 36 and 45 degrees to align with flat plate normals of target class 2 and 3, respectively. This radar platform arrangement supports the development of a monostatic ATR performance baseline. The performance baseline is used for non-coherent gain analysis and multi-platform, multi-static ATR performance comparisons.

Table 5. Classifier test summary for test scenario 1 including platform configurations and a brief test description.

Test Number	Platform Configuration(s)	Test Description
1	1,2, and 4	Multi-static classifier performance improvements are not solely the result of non-coherent gains
2	4	Multi-static classifier performance increases with an increasing number of dominant scatterers
3	1, 2, 3, and 4	Multi-static platform configurations outperform monostatic platform configurations
4	4	Multi-static classifier performance varies based upon placement of participating platforms

Platform configurations 3 and 4 rely on the geometry shown in Figure 14 to show the benefits of multi-static reflectivity. With platforms positioned at 0, 36, and 90 degrees, multi-static reflectivity contributions are observed at each platform according to Table 3. The multi-static reflectivity contributions enable the classifier to demonstrate how dominant scatterers and platform placement influence classifier behavior.

4.2.3.2 Test Scenario 2

A second set of tests were conducted to show the classifier's robustness to target pose angle uncertainty. Three target pose uncertainty models (known, Gaussian, and uniform) were

implemented for the actual test target data and classifier to examine performance relative to test and training data model mismatches. The tests executed in the second scenario are provided in Table 6.

Several Monte Carlo simulations were performed using test data with actual target pose angle distributions following one of the three target pose uncertainty models (known, Gaussian, and uniform). Each of the three target pose uncertainty models, in turn, was assumed by the classifier in these simulations. These simulations support the analysis of nine different combinations of the pose model used to generate the test data and the pose model assumed by the classifier. A second set of simulations were executed to compare classifier performance when both the actual target pose angle and classifier assume target pose certainty is governed by a Gaussian distribution. These simulations compare different Gaussian variance combinations for actual data and assumed classifier target pose to evaluate classification sensitivity to target pose uncertainty.

Table 6. Classifier test summary for test scenario 2 including platform configurations and a brief test description.

Test Number	Platform Configuration(s)	Test Description
5	4	Multi-static classifier target pose angle assumptions affect classification performance for known actual target pose angle
6	4	Multi-static classifier performance is sensitive to matching variance parameters for actual and classifier Gaussian target pose angle models
7	4	Multi-static classifier target pose angle assumptions affect classification performance for uniform actual target pose angle

4.2.3.3 Test Scenario 3

The third test scenario assesses the effects of CCD on the proposed multi-static, distributed, multi-sensor ATR solution. Under this scenario, target class 1 test data was modified to mimic three different CCD attempts. The description and purpose for each target class 1 configuration is detailed in Table 7. Test number 8 was implemented by placing a single, spheroidal scatterer with an RCS of 30 dBsm in the region spanned by the points [0.5, 0.0] and [1.0, 0.5] in Figure 12(a). This scatterer mimicked the behavior of a cylindrical pole attached to the target extending in the z - coordinate direction. Test number 9 required uniformly modified RCS values for non-dominant scatterers located in the region spanned by the points [0.0, 0.5] and [0.5, 1.0] in Figure 12(a). Uniform modification of the RCS values in this region simulates the camouflage of the dominant scatterer located within the region. Finally, test number 10 artificially adjusts two of the four flat plates attributable to target class 1. The normals associated with these two plates were modified to 36 degrees in an attempt to confuse the classifier into thinking the modified target class 1 was similar to target class 2.

Table 7. Classifier test summary for test scenario 3 including platform configurations and a brief test description.

Test Number	Platform Configuration(s)	Test Description
8	4	Multi-static classifier performance for target with unknown, spheroid dominant scatterer
9	4	Multi-static classifier performance for target with unknown, regional dominant scatterers
10	4	Multi-static classifier performance for target with unknown, modification of dominant scatterer plate normals

4.2.4 Test Procedures

Simulated HRR profile data sets were generated without additive noise and stored for repetitive use during all tests. Parameterized AWGN was added to the raw HRR profiles to generate simulated test data. Once the appropriate AWGN was added to the raw HRR profile data, each profile was subjected to windowing and normalization. The windowing operation simply removed range bins that did not include target returns. For the targets described in Section 4.2.2, the window was applied to all HRR profiles permitting contributions from range bins 9 through 24. The normalization process used the L_2-norm to equalize the power of the HRR profiles. This equalization ensured the classifier was discriminating target classes based upon spatial signature differences and not differences in total signal power. Similar to the argument presented in Section 3.2 regarding the inability to use phase in the classifier, normalization of the HRR profiles reduces the impact of differences in signal power between the test and training data sets on classifier operation.

Requisite test statistics were used as inputs to a forced decision classifier. The classifier was executed for each of the 45 slow-time samples within the test HRR profile data set. A conservative approach was instituted regarding noise in the test and training data sets. For each test sample, the classifier used a training data set with the same noise level as the test data set. In some studies, researchers use "ideal" training data containing little to no noise. The conservative approach used in this study challenges the classification with respect to imperfect training data. Even if the training data is "collected" under seemingly pristine conditions, there may be systematic errors, such as mismatches between the scale model used in the compact radar range collection or the CAD model employed by the EM code and the real target; the noise in the training data is intended to encapsulate these effects. In practice, test data will typically be

collected under noisier conditions than the training data; future work could include studying the effects of varying noise in the training data relative to the test data. In addition to allowing us to stress the overall ATR approach, adding noise to the training data helps us avoid analytic complications in statistically modeling the magnitude data, as described in Section 3.2. If we do not wish to stress the algorithm in terms of limited training data, we can simply increase the number of Monte Carlo runs used to generate the training set.

Tests were executed at 10 different noise power levels for each of the three target classes. The classifier decisions were tallied for each test and training target class combination to produce a confusion matrix. A percentage of correct classification was calculated for each cell within the confusion matrix. Unless otherwise noted, the results presented in Section 4.3 use mean PCC as the primary performance criterion.

Several of the curves presented in Section 4.3 show somewhat erratic behavior at the high noise power indices. This behavior is attributable to a thresholding effect [39] on the classifier caused by significant noise spikes; we expect that additional Monte Carlo runs would smooth out these curves. This behavior does not diminish the general conclusions of these results since:

1. Classification results with PCC levels below 70% are unlikely to be considered relevant given challenge #3 in Chapter 1;
2. The trends observed in other sections of the curves are consistent with expected behavior; and,
3. Many of the percentages of correct classification in these regions are close to the statistical coin flip percentage of 33%, indicating a noise level at which the system would be unusable anyway.

4.3 Simulation Results

The three test scenarios described in Section 4.2.3 were executed using the data and signal models described previously. The results of each test scenario are detailed in the sections below.

4.3.1 Test Scenario 1

The simulations associated with test scenario 1 were designed to demonstrate several baseline performance measurements. In particular, we would like to ensure that performance improvements are attributable to spatial diversity and not simply to the non-coherent gain benefits of multiple classifier inputs. Non-coherent summation of data can increase SNR [40], and any increase in SNR has the opportunity to improve ATR performance. Thus, it is important to qualify the impact of any non-coherent gain when incorporating data from more than one platform.

For the spatial diversity versus non-coherent gain analysis, each radar platform was placed to leverage the dominant scatterer reflectivity of each target class. This platform configuration presumably offers the best opportunity to distinguish between each target class when all three radar platforms are used in (4.8). To demonstrate ATR performance improvement due to non-coherent summation of test statistics, nine independent test runs were executed with the stated platform configuration. The classification results are presented in Figure 15. The solid blue (lowest) line represents the mean PCC for the set of slow-time samples collected at platform 1. The red (second lowest) and green (second highest) lines denote the classifier's performance when the test statistics from three and nine independent slow-time sample collections at platform 1 are summed and submitted to the classifier for processing, respectively. As can be seen, there is a quantifiable improvement in platform 1 mean PCC when more than one sample is used for classification.

However, multiple, spatially-diverse sensors are capable of significantly outperforming an equal number of single-sensor samples. This performance improvement is noticeable in Figure 15 by observing the solid magenta (highest) line. With the exception of the highest noise level, the inclusion of multiple, spatially-diverse platforms resulted in a mean PCC that exceeded the performance of the multiple-sample, single-platform configuration by 9 dB or more. This conclusion demonstrates the advantages of employing multiple platforms at different target pose angles, as indicated in [15].

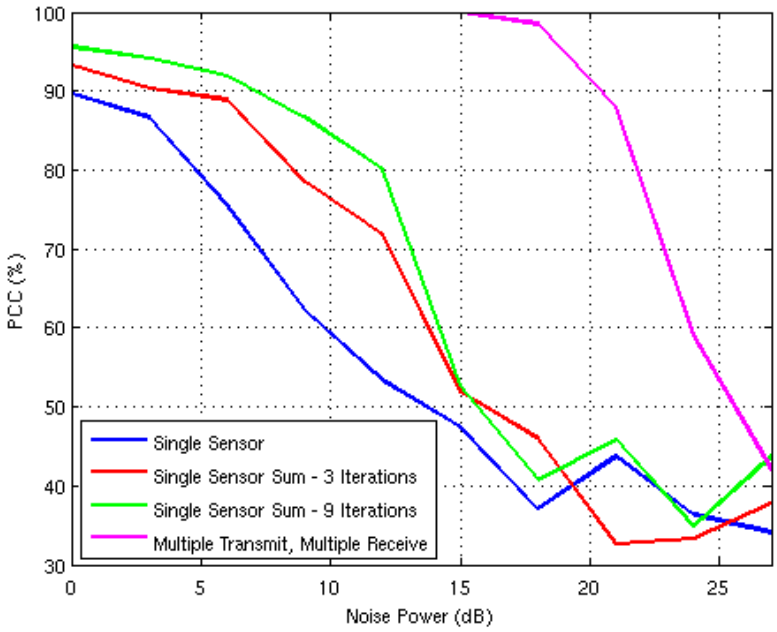


Figure 15. Performance comparison of single sensor, summed single sensor, and multi-sensor classification.

The next set of simulations for test scenario 1 addressed the role of dominant scatterers in classifier performance. This set of simulations employed platform configuration 4, with platforms positioned at 0, 36, and 90 degrees. The positions of the platforms provide the

classifier with potentially more discernible views of each target class, as defined in Table 3. Target class identification by transmit and receive platformThe ability to incorporate multi-static reflectivity at a specific platform allows the classification algorithm to more easily distinguish between target classes with different dominant scatterer characteristics. This is evident by evaluating the performance of the classifier while varying the number of dominant plate scatterers included within the target class. The potential benefits of multi-static reflectivity from dominant scatterers are illustrated by observing mean PCC in Figure 16. Classification performance is significantly lower in the instance where each target class lacks dominant scatterers. As more dominant scatterers are included in each target class, the plot shows better performance at the higher noise levels.

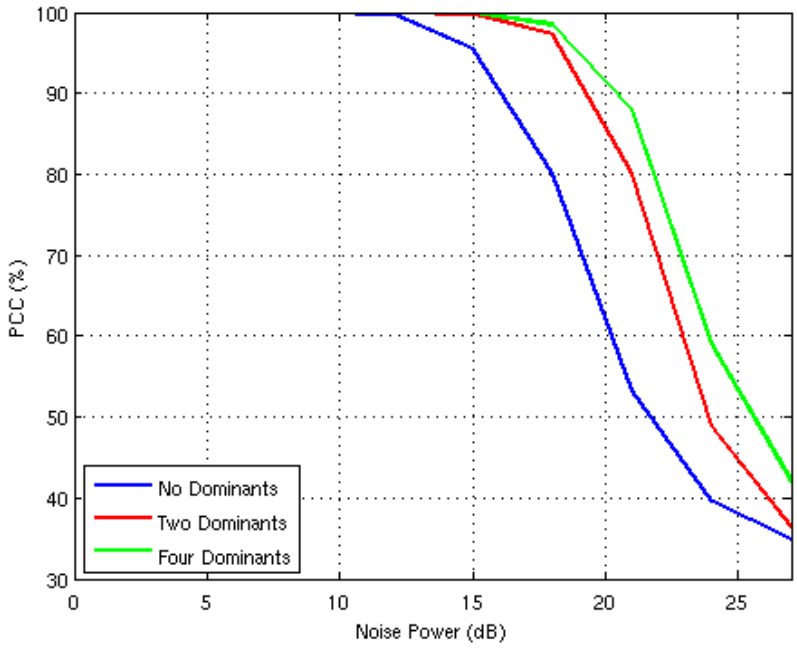


Figure 16. Mean PCC performance demonstrating the benefits of having more dominant scatterers.

The next step is to evaluate PCC performance relative to different measurement scenarios. A third set of simulations was run to examine mean PCC performance for each platform configuration outlined in Table 4. The results of these tests are depicted in Figure 17. It is clear from these results that more than one sensor provides improved PCC when compared to a single sensor. The improved PCC performance measures for multiple, spatially-diverse views are consistent with the results achieved by previous studies [18, 11, 15, 19, 41]. The blue (lowest) curve in Figure 17 represents the best mean PCC across all three independent, monostatic platforms. The remaining curves show mean PCC performance for multi-sensor configurations. In each multi-sensor configuration, the mean PCC curve exceeds the best monostatic curve, illustrating the potential advantage of multiple sensors. The benefits of exploiting multi-static reflectivity are demonstrated by the superior performance of the multiple-transmit, multiple-receive platform configuration (magenta, highest curve) compared to the summed monostatic platform configuration (red, second lowest curve). The potential to improve classifier performance when multiple platforms transmit and receive versus just receive is evident from the comparison of the magenta (highest) and green (second highest) curves.

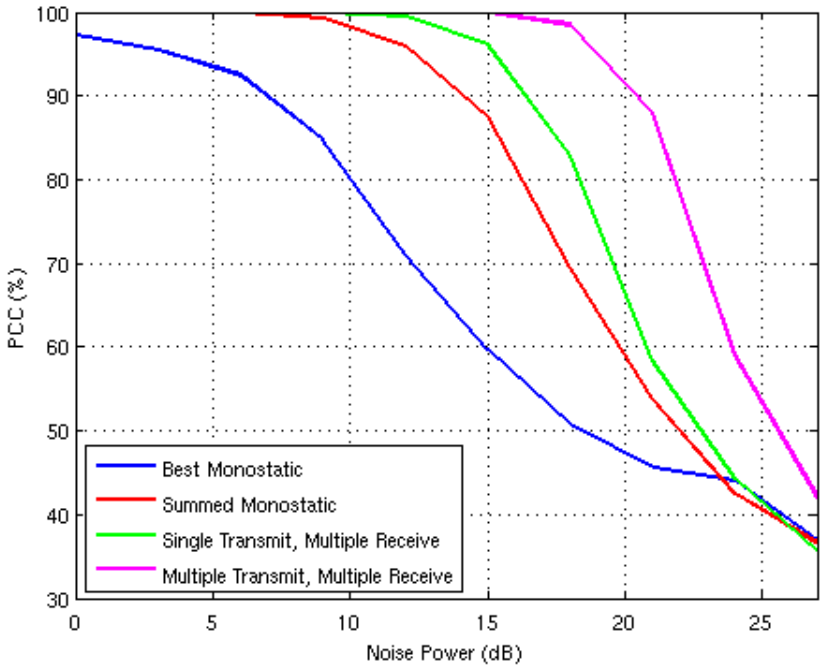


Figure 17. Mean PCC performance showing relative performance of the four platform configurations.

The multi-platform performance results in Figure 17 are encouraging, but not indicative of all possible sensor locations. In Figure 18, the mean PCC performance for the multiple-transmit, multiple-receive platform configuration is shown for two different sensor placement combinations. Three platforms positioned at 0, 36, and 90 degrees provide good performance, as demonstrated by the blue (highest) curve. However, positioning the three platforms at 81, 153, and 159 degrees results in a significant reduction in mean PCC, as noted by the red (lowest) curve. Thus, a better understanding of where to position the platforms would be helpful in improving performance. The examination of the relationship between target RCS, multi-sensor placement, and PCC is the subject of Chapter 5.

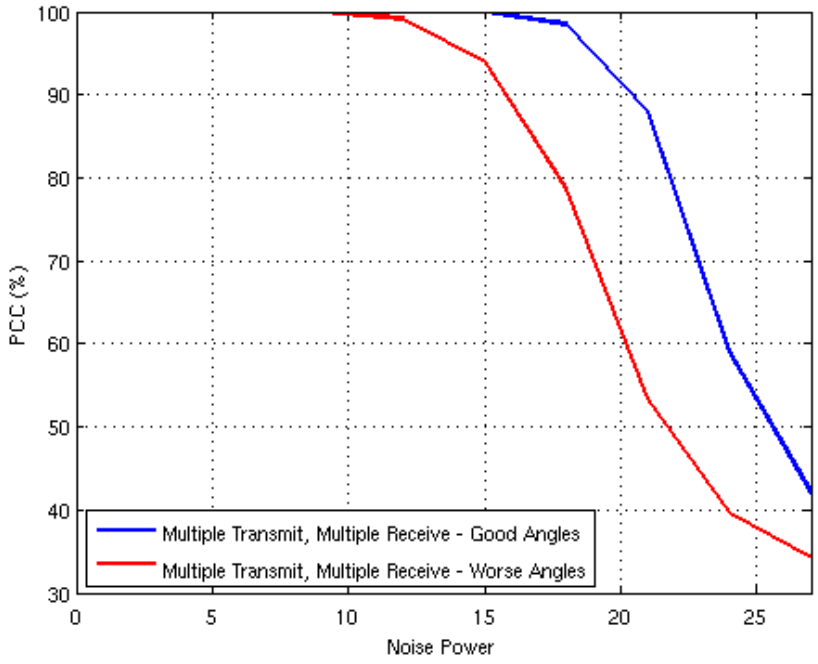


Figure 18. Mean PCC performance demonstrating sensor placement impacts classifier performance.

4.3.2 Test Scenario 2

Test scenario 2 simulations addressed the sensitivity of matching the classifier's assumed prior distributions on pose to the poses in actual test data. The first series of Monte Carlo simulations in test scenario 2 evaluated classifier performance under the guise of a known target pose angle for the test data. In these simulations, test data from a single target pose angle is fed to the classifier. Each of the three target pose uncertainty models (known, Gaussian, and uniform) was used by the classifier to produce the results shown in Figure 19. The known "true" target pose angle results exhibited expected behavior. The known target pose angle uncertainty model for the classifier provided the best results; it represents an upper bound on classifier performance. The lower bound on classifier performance was achieved by using the uniform target pose angle

uncertainty model for the classifier. Using the Gaussian model for target pose angle uncertainty in the classifier produced classification rates within the upper and lower bounds. The Gaussian uncertainty model also demonstrated expected behavior as higher variances are used in the model. Note the progressively poorer performance as the variance is increased from 1 to 100 angle indices.

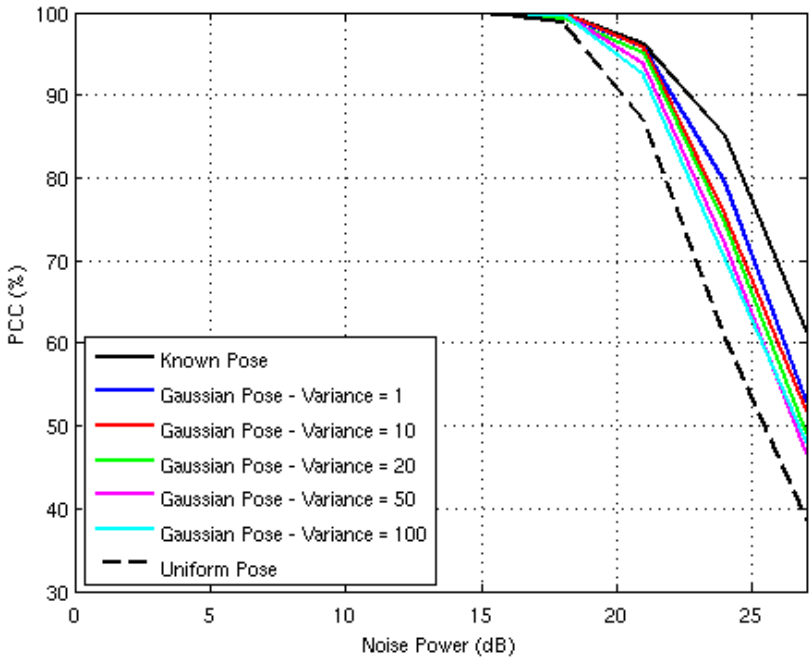
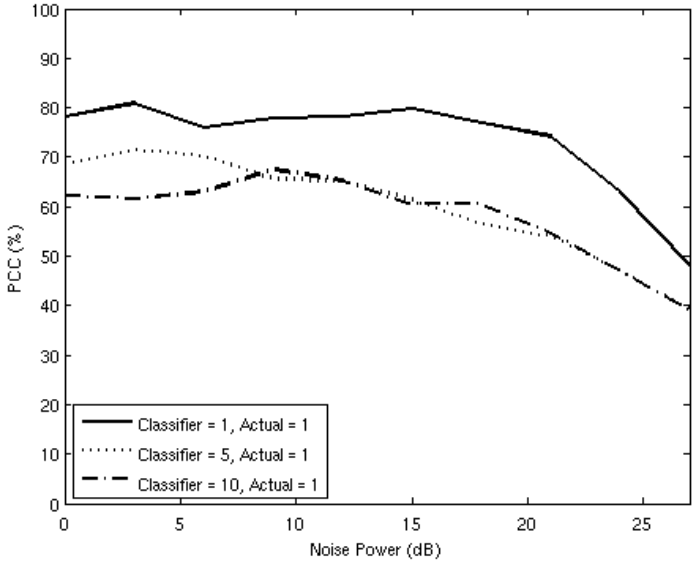


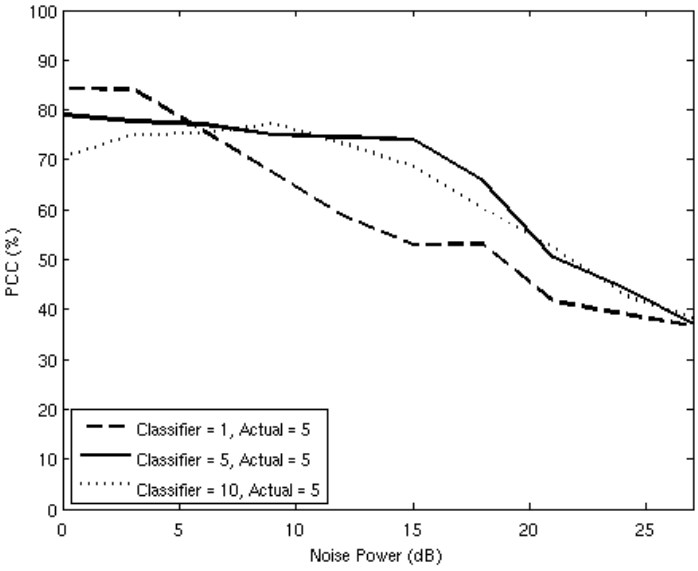
Figure 19. Comparison of classifier results showing performance differences of each classifier pose distribution assumption when actual pose is fixed.

The second series of simulations in test scenario 2 compared PCC results for different combinations of test and classifier-assumed Gaussian target pose uncertainty model standard deviations. The results for these simulations are shown in Figure 21. In each test, better PCC performance was obtained when the classifiers' standard deviation assumption matched the actual test data standard deviation. The results are marginally convincing for a standard deviation of

five, as noted in Figure 21 (b). However, the helpfulness of matching the classifiers' assumed pose model to that of the actual test data is evident by examining Figure 21 (a) and Figure 21 (c).

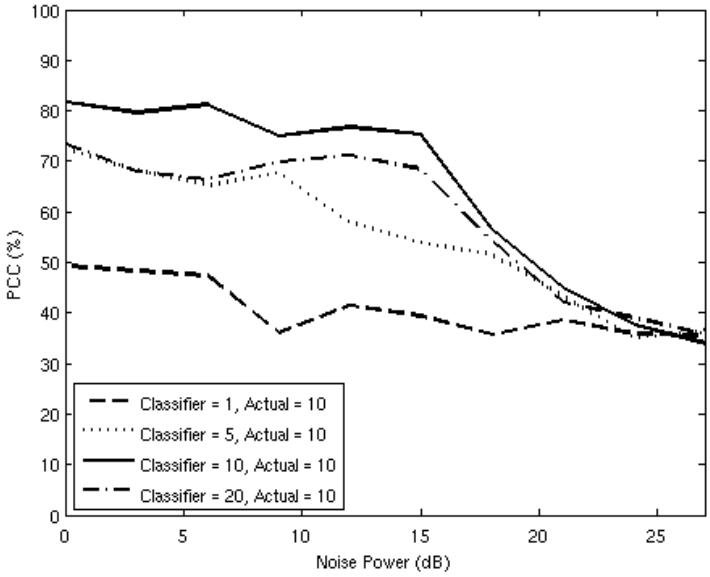


(a)



(b)

Figure 20. PCC comparison of Gaussian target pose uncertainty models for actual test data and the classifier: a) actual standard deviation = 1, b) actual standard deviation = 5, and actual standard deviation = 10.



(c)

Figure 21. Continued.

When knowledge of the actual target pose angle standard deviation is less certain, the results of Figure 21 (c) indicate it is better to assume a wider standard deviation for the classifier. A narrower assumption on target pose angle uncertainty for the classifier produces poor results when the actual target pose angle uncertainty is wide (dashed curve). However, a broader assumption on target pose angle uncertainty gives more favorable results (dash-dot curve).

The third series of simulations examined classifier performance using data with a pose angle drawn uniformly. Classifier performance under these conditions is shown in Figure 22. As expected, the uniform target pose angle uncertainty performed the best, considering the narrower certainty assumptions inherent in the known and Gaussian target pose angle models. The

assumed "known" target pose angle performance mimics selection based upon a three-sided coin flip. In this case, the PCC values hover around 33% for all noise levels.

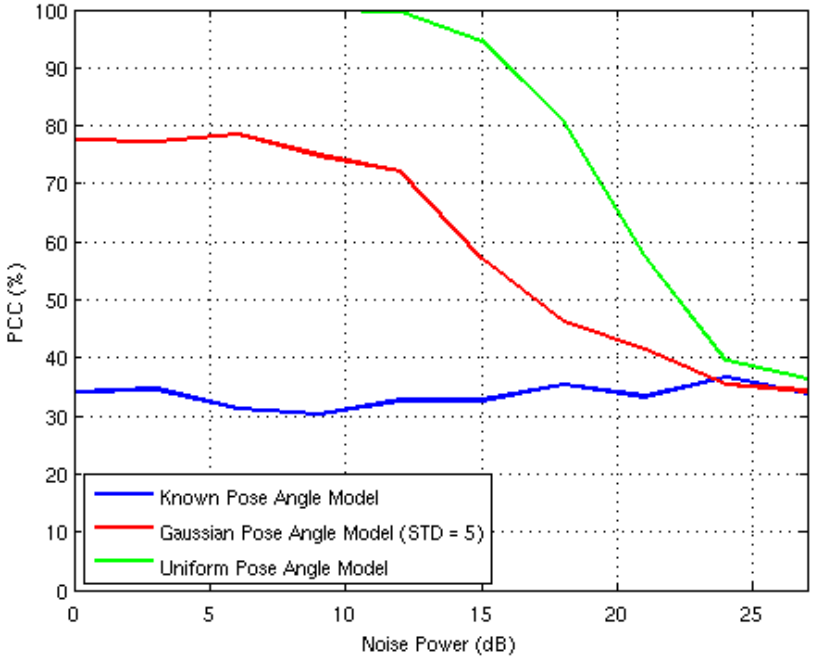


Figure 22. Classifier performance for actual target pose angle uncertainty modeled as uniform.

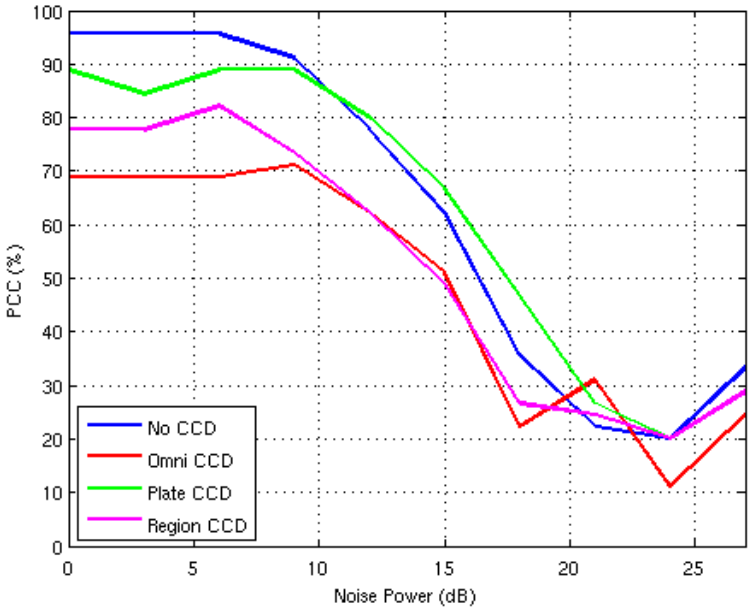
Note that in general, as the test data target pose angle becomes more uncertain, the classifier performance decreases. For example, comparison of the uniform target pose angle results from Figure 19 and Figure 22 show degrading PCC performance as the test data target pose angle becomes less certain.

4.3.3 Test Scenario 3

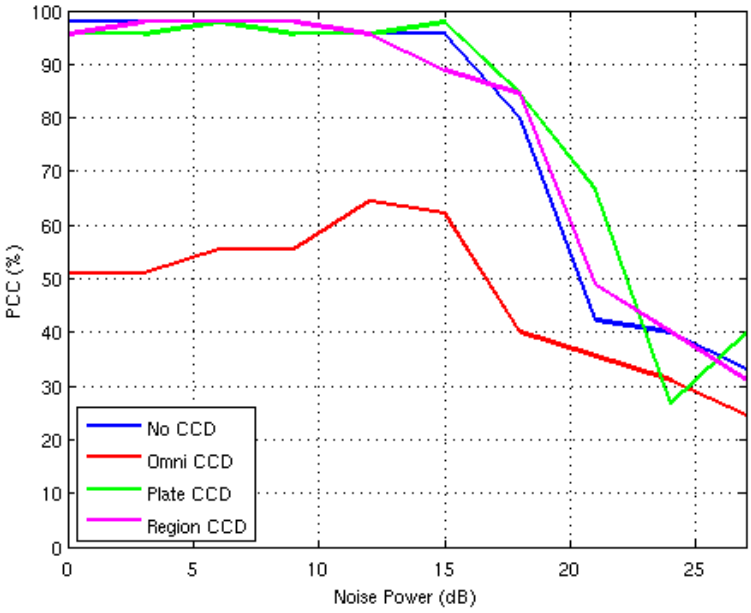
The goal of test scenario 3 was to evaluate the impact of various CCD effects on the derived classifier. The intent of tests 8, 9, and 10 in Table 7 was to confuse the classifier by introducing

an omnidirectional scatterer, camouflaging a dominant scatterer, or reconfiguring the plate normals of target class 1, respectively. Several simulations were executed with the same sensor placement – 0, 36, and 90 degrees – that produced the results presented in Figure 17. Test data was generated for each of the test target configurations using the previously generated target class 1 data set as a base. Each test target configuration was tested independently to produce confusion matrices that included all three target classes. The resulting confusion matrices were used to derive PCC results related to tests against the non-CCD version of target class 1. Thus, the PCC results explore the robustness of the classifier, relative to RCS modifications mimicking target class 1. PCC results for each of the test target configurations using a monostatic configuration are shown in Figure 24. These results can be contrasted with the multiple-transmit, multiple-receive platform configuration classifier output shown in Figure 25.

The first key observation from Figure 24 is that all platforms show degradation in target classification to varying degrees for each of the CCD modifications. These results are expected since the training data does not include the tested modifications, so the classifier has more difficulty in matching the test data to the training data. The second key observation is how performance varies based upon platform placement. When compared to the no-CCD performance curves, platforms 2 (36 degrees) and 3 (90 degrees) do a reasonable job in properly classifying target class 1 with plate and region modifications. With respect to the omnidirectional scatterer, platforms 1 (0 degrees) and 2 demonstrate a significant inability to properly classify the tested target as target class 1.

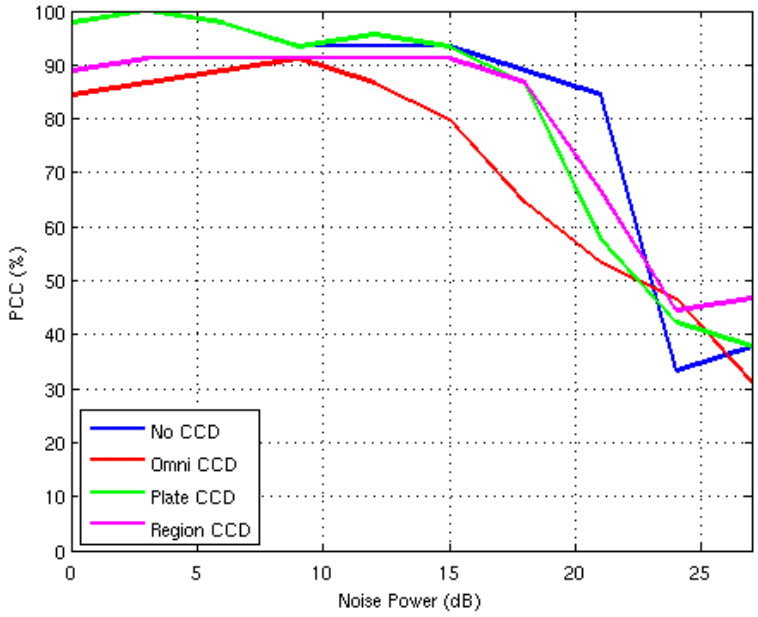


(a)



(b)

Figure 23. Monostatic PCC performance for target class 1 when subjected to CCD effects: a) platform 1 (top), b) platform 2 (middle), and c) platform 3 (bottom).



(c)

Figure 24. Continued.

The variable and poor performance attributable to each platform in a monostatic configuration is improved dramatically when multi-static returns are incorporated. The classification results of the multiple-transmit, multiple-receive platform configuration are provided in Figure 25. One of the most significant improvements is achieving ninety percent or better PCC up to and including 18 dB noise power regardless of the CCD modification. Even the most effective CCD technique in this platform configuration – plate modification – has a PCC curve that meets or exceeds the best curve in Figure 24. Another important observation from Figure 25 is how well the multiple-transmit, multiple-platform configuration performed for each of the CCD modifications when compared to no-CCD modification. The classification curves for each of the CCD modifications mimic the no-CCD curve with a reduction in PCC of no worse

than ten percent. Interestingly, there appears to be no benefit to using a regional CCD effect since its PCC curve is almost identical to the no-CCD curve.

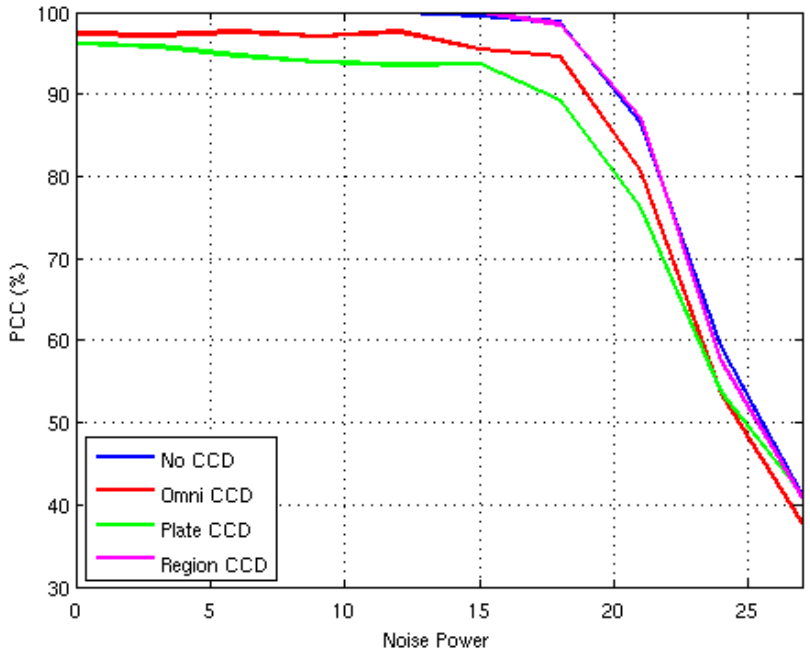


Figure 25. Multiple-transmit / multiple-receive PCC performance for target class 1 when subjected to CCD effects.

4.4 Summary

A multi-static, distributed, multi-sensor classifier was derived and shown to be able to perform better than a monostatic, distributed, multi-sensor classifier. The performance improvements attributable to the new classifier were achieved by leveraging multi-static returns. Dominant scatterer returns facilitated better classifier performance as the number of dominant scatterers increased. Further, our simulations illustrated that the performance increases arising from using multi-static data were not merely due to non-coherent gain improving the signal-to-noise ratio; these improvements arose from the way scattering varies with angle.

The derived classifier was studied under conditions of varying target pose angle uncertainty. Simulations using a known target pose angle demonstrated the upper performance bound for the classifier. Gaussian target pose uncertainty models for test data generation and the classifier were explored in varying combinations to demonstrate levels of performance. The principal conclusion from these studies is that if there is any doubt about how accurately the pose of the target is known, it is generally wiser to be pessimistic and assume a larger standard deviation. Guessing "wider" does not accrue the same kind of penalty as guessing "narrower." An unjustifiably overconfident classifier will tend to deceive itself.

The impact of CCD effects on the derived classifier was also explored. Target class 1 was modified to exhibit three different CCD effects: 1) single omni-directional scatterer, 2) camouflage region covering a dominant scatterer, and 3) modification of dominant scatterer plate normals intended to mimic target class 2. Simulations were performed to examine how each CCD effect influenced monostatic classifier performance at each radar platform. The results demonstrated varying levels of confusion at each platform. Marginal monostatic classifier performance was compared to classification rates achieved with a multi-static, distributed, mulit-sensor platform configuration. The multi-static, distributed, multi-sensor classifier demonstrated exceptional performance with regard to ignoring the CCD effects. The maximum deviation in classifier performance between targets with and without CCD effects was 10%.

Finally, the derived classifier's dependence on sensor placement was examined. Two different combinations of sensor placement were used to demonstrate significant differences in classifier performance. Classification performance reductions of 30% were noted for some noise power levels. Similarly, a reduction of 5 dB in noise power was required to maintain a reasonable classification rate in some simulations. These results indicate a substantial need for

understanding the role of sensor placement in the multi-static, distributed, multi-sensor classifier. This topic is explored in the next section, Chapter 5.

5 USING AN INFORMATION THEORETIC-BASED DISTANCE MEASURE TO IMPROVE SENSOR PLACEMENT

This chapter details an approach for identifying sensor placements with a potential for providing high target classification rates. The approach outlined uses a surrogate measure, Kullback-Liebler divergence (KLD), in place of probability of error to establish a relationship between differences in probable target class probability distributions and PCC. Due to some strict conditions required by the KLD derivation that cannot be met, in general, by our problem, a lower bound on an approximate KLD is derived. Henceforth, the lower bound on the approximate KLD is referenced as "the KLD" for brevity.

The KLD is parameterized by relative transmit incidence and receiver angles to support comparisons of KLD and PCC values for various sensor placements. The KLD and PCC analyses are augmented by the development of two sensor placement algorithms. An optimal algorithm is derived that leverages an exhaustive search of all possible transmitter and receiver angles. A more efficient approximate algorithm is also developed to analyze the tradeoffs between reduced algorithmic complexity and classification rates. Simulations are conducted to demonstrate the benefits of using either algorithm for improved target classification.

5.1 Sensor Placement Algorithms

5.1.1 Distance Measure

The notion of a distance measure is pertinent to identifying optimal sensor placement within a multi-sensor ATR environment. Since the HRR profiles obtained by each sensor are stochastic, target classification depends upon the probability densities describing the HRR profiles. The characteristics of these distributions dictate the probability of misclassification.

The goal of this chapter is to find configurations where the classifier identifies the probable target class with a minimal probability of error. Minimizing the probability of error with multi-dimensional data sets is often difficult. In such cases, a reasonable substitute for the probability of error may be a distance measure between the probability distributions of competing target classes.

There is a rich literature relating to probability of error and various distance measures, such as Kullback-Leibler, J-divergence, and Bhattacharyya measures. In [42], the use of a suboptimal distance measure was shown to be an effective surrogate for the minimum probability of error. In this dissertation, the KLD examined in [42] has been selected to assist with sensor placement. The derivation of the KLD in the context of a multi-static, distributed, multi-sensor environment is provided in the next section.

5.1.2 Kullback-Leibler Divergence

The concept of a distance measure between probability densities is rooted in information theory. In [43], Kullback and Leibler established a measure for discriminating between two probabilistic hypotheses[5]. KLD is defined as the "mean information for discrimination between H_i and H_j," where H_i denotes the hypothesis of target class i being the true target class. The term "information" refers to the classical definition as provided in [44] and introduced by Shannon in [45]. Information can be described as the amount of uncertainty associated with a random variable. Because the KLD characterizes the separation between two hypotheses, the relevant random variable must be a function of the two hypotheses' probabilistic representations.

[5] The Kullback-Leibler distance measure does not satisfy the triangle inequality due to its asymmetry. However, most literature still refers to the resultant as the Kullback-Leibler distance measure. We use the same liberties throughout this thesis.

The random variable provided in [43] is prescribed as the ratio of the two applicable probability densities,

$$I\left(\hat{\mathbf{z}}\right)=\ln\frac{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_j;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}\,, \tag{5.1}$$

where $\hat{\mathbf{z}}_{a,b}\in\hat{\mathbf{Z}}$. The mean information is calculated with respect to the probability distribution associated with hypothesis i,

$$E_i\left[I\left(\hat{\mathbf{z}}\right)\right]=\int_{Z_i}p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)\ln\frac{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_j;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}d\hat{\mathbf{z}}\,, \tag{5.2}$$

where the dependence of $\hat{\mathbf{z}}$ on transmit, receive, and target pose angles has been suppressed for clarity; and, Z_i designates the support set relevant to the expectation operation, assured to include all sample space elements for both densities.

The multi-dimensional integral defined in (5.2) requires integration over all possible target pose angles. For purposes of simplifying the conceptual and computational requirements of this study, the derivation of the KLD employs a chosen "true" target pose angle for the referential target class i, $\theta_{pose}^{'}=\theta_{pose}^{true}$ although the classifier is unaware of this "true" pose angle. The detailed derivation of (5.2) is provided in the Appendix.

The derived KLD provides a single scalar representing the difference in the multivariate probability distributions. Our sensor placement algorithms seek to identify the transmit and receive angle combinations producing the largest KLD value, so we modify (A.17) to explicitly

express KLD as a function of $\boldsymbol{\theta}_{a,b}$. Removing the summation operators reveals the desired target class divergence measure,

$$KLD\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)=\underbrace{-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)}_{5.3.1}-\underbrace{P}_{5.3.2}-\max_{\theta_{pose}}\left\{\underbrace{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)}_{5.3.3}\right.$$
$$-tr\left[\underbrace{C_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)}_{5.3.4}\right] \qquad (5.3)$$
$$\left.+\underbrace{\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)}_{5.3.5}\right\}.$$

Equation (5.3) represents the base expression used to develop the optimal and approximate sensor placement algorithms in Sections 5.1.4 and 5.1.5, respectively. A deeper understanding of the KLD measure can be obtained by examining contributions from each of the terms in (5.3). Terms 5.3.1 and 5.3.3 are scaling factors to encapsulate the effects of target power. This is evident by analyzing the covariance determinant operation. Since the determinant of the covariance matrix is the product of the eigenvalues, which are representative of the power in each HRR profile dimension, the determinant operation provides a general measure of power in the target return. Since the covariance matrices are derived from normalized HRR profiles, Terms 5.3.2 and 5.3.4 provide a relative distance measure based upon covariance structure. If C_i and C_j have the same structure, terms 5.3.2 and 5.3.4 cancel one another. However, if there is a difference in structure between C_i and C_j, the KLD value is increased. Finally, term 5.3.5 is the Mahalanobis distance measure for $\boldsymbol{\mu}_i$ in relation to the probability distribution statistics associated with target class j. This term provides a general measure of how different the mean HRR profile for target class i is when compared to the statistical properties of target class j.

When interpreting these formulas, remember that the data used is the magnitude of the HRR profiles, and not the original complex values.

5.1.3 Operational Scenario Definition

Consider a scenario where three radar platforms are coordinating positions relative to a ground-based, stationary target. One radar platform is responsible for transmission, while the other two platforms are receivers. Sensor data from the two receive platforms is transmitted to a central radar platform. This central processing platform could be, but is not necessarily, the transmission platform. An example scenario is provided in Figure 26.

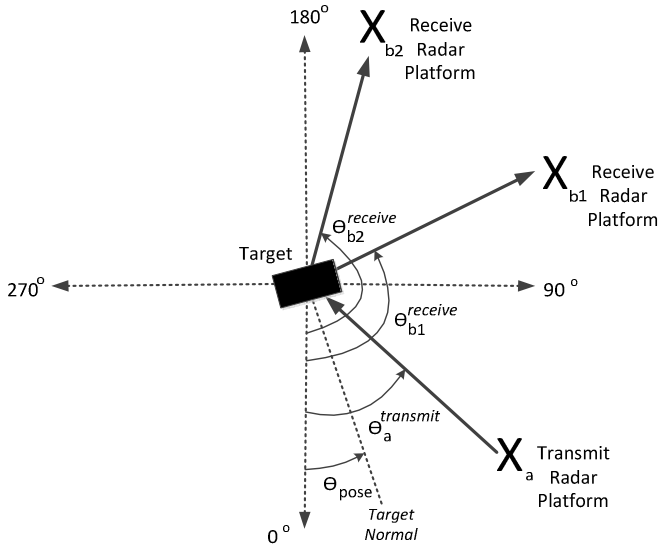


Figure 26. Scene geometry demonstrating multi-platform interrogation of a ground-based, stationary target.

Each receiver (X_{b1} and X_{b2}) is responsible for collecting reflectivity data within a prescribed slow-time data collection time span. Data is captured via a straight line flight path tangent to a consistent target range radius, as described in Section 4.2.1. HRR profiles are

generated from the target reflectivity data and sent to a central processing platform employing the classifier defined in Section 4.1.

5.1.4 Optimal Sensor Placement Algorithm

The premise for the optimal sensor placement algorithm is to identify the best placement of a single transmitter and two receive platforms as determined by the highest KLD value between two target class probability densities. A three-dimensional placement algorithm can be achieved by parameterizing (A.17) in accordance with the three angles shown in Figure 26,

$$\begin{aligned} KLD\left(\theta_a,\theta_{b1},\theta_{b2},\theta_{pose}^{true}\right) = & -\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{true}\right)\right)\right)-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{true}\right)\right)\right)-P \\ & +\max_{\theta_{pose}^{"}}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{"}\right)\right)\right)-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{"}\right)\right)\right)\right. \\ & -tr\left[C_i\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{true}\right)\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{"}\right)\right] \\ & +\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{"}\right)\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b1},\theta_{pose}^{"}\right)\right) \\ & -tr\left[C_i\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{true}\right)\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{"}\right)\right] \\ & \left.+\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{"}\right)\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b2},\theta_{pose}^{"}\right)\right)\right\}. \end{aligned} \quad (5.4)$$

Equation (5.4) is used to identify the optimal transmitter and receiver placements according to

$$\left[\hat{\theta}_a,\hat{\theta}_{b1},\hat{\theta}_{b2}\right]_{optimal} = \arg\max_{\theta_a,\theta_{b1},\theta_{b2}}\left\{KLD\left(\theta_a,\theta_{b1},\theta_{b2},\theta_{pose}^{true}\right)\right\}, \quad (5.5)$$

where $\hat{\theta}_a$, $\hat{\theta}_{b1}$, and $\hat{\theta}_{b2}$ are chosen placements of the transmitter and receivers, respectively.

5.1.5 Sensor Placement Approximation Algorithm

The optimal sensor placement algorithm specified in Section 5.1.4 requires a KLD computation for each permutation of the three independent reference angles. Although this will typically be offline computation, the processing requirements could prove burdensome in some instances particularly if the algorithm was extended to include additional platforms. A less

computationally-intensive approach could begin by using (5.3) to find the transmit and receive angles for a two-platform scenario. Then, using the specified transmit angle and receive angle, the algorithm examines KLD values for all possible receive angles to select a third platform's receive angle. This approximate sensor placement approximation algorithm can be formalized as

$$\text{Stage 1: } \left[\hat{\theta}_a, \hat{\theta}_{b1}\right] = \arg\max_{\theta_a, \theta_{b1}} \left\{ KLD\left(\theta_a, \theta_{b1}, \theta_{pose}^{true}\right)\right\}$$

$$\text{Stage 2: } \left[\hat{\theta}_a, \hat{\theta}_{b1}, \hat{\theta}_{b2}\right]_{approx} = \arg\max_{\theta_{b2}} \left\{ KLD\left(\hat{\theta}_a, \hat{\theta}_{b1}, \theta_{b2}, \theta_{pose}^{true}\right)\right\}, \quad \theta_{b2} \neq \hat{\theta}_{b1} \tag{5.6}$$

One improvement on this approximate algorithm might be to add additional steps cycling through the parameters; for instance, we could add a Stage 3 in which we try to refine θ_a by maximizing over it while keeping the receive angles found in Stages 1 and 2 fixed, and so on. More complicated scenarios with more than three platforms might benefit from extension of such approaches; for instance, a scenario with six platforms might involve optimizing each parameter in turn while holding the other four fixed, cycling through in various combinations. We leave these issues as avenues for future work.

5.2 Simulation

The effectiveness of using KLD to predict classifier performance and place platforms is examined using KLD sets for a two-class scenario. The two target classes exhibit significant similarities at many angles and dissimilarities at other angles, as described in Section 4.2.2.1. The algorithms specified in (5.5) and (5.6) require multi-static HRR profile data for two target classes at each transmit (θ_a) and receive (θ_b) angle combination. Using the operational scenario defined in Section 5.1.3, bistatic HRR profile data was simulated according to the specifications in Chapter 3. Multi-static HRR profiles generated using the prescribed straight-line flight path

produce 90 slow-time samples per receive angle. Odd-numbered slow-time samples form a 45-sample data set for each transmit and receive angle combination. Additive, white Gaussian noise was added to each bistatic HRR profile for thirteen noise levels in 3 dB increments, 0 dB to 36 dB. These data sets were used by each angle selection algorithm to identify sensor placements for each noise level. All simulations were executed using a true target pose (θ_{pose}^{true}) of zero degrees.

5.3 Simulation Results

5.3.1 Single Transmitter, Single Receiver Results

5.3.1.1 KLD Surface Plots

The KLD values generated by (5.3) produce a surface parameterized by transmit and receive angles. This surface provides a clear indication of where two target classes are similar and dissimilar. The KLD values plotted in Figure 27 were computed using (5.3) under an expectation operation with respect to target class 1.

Several key observations are evident from Figure 27. First, the dominant scatterers attributable to target class 1 result in relatively large KLD values, as shown by the diagonal line traversing the plot from upper-right to lower left. These lines correspond to the distinctive flat plate returns from target class 1 as the transmitting platform transitions through all 120 possible transmit angles. The returns from target class 2's flat plates are not noticeable in Figure 27. However, returns from target class 2 flat plates are apparent when the expectation operation in (5.3) is performed with respect to the target class 2 probability distribution. The KLD surface of class 2 versus target class 1 taken with respect to the target class 2 probability distribution is shown in Figure 28.

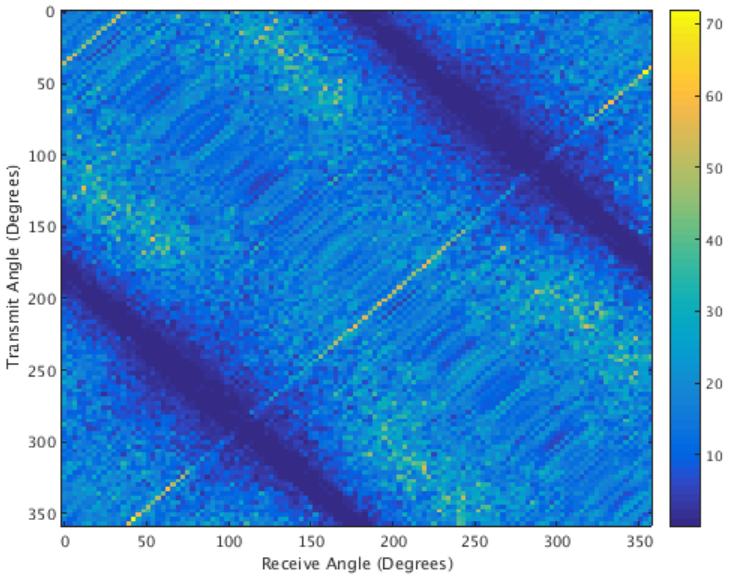


Figure 27. Surface plot showing KLD values comparing target classes 1 and 2 at 9 dB noise power.

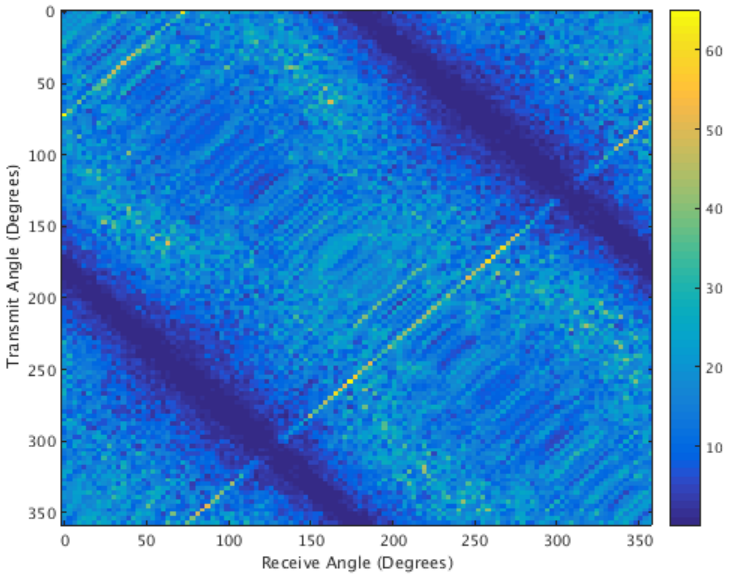


Figure 28. Surface plot showing KLD values comparing target classes 2 and 1 at 9 dB noise power.

In Figure 28, the distinctive diagonal lines identifying flat plate scatterer returns have shifted in accordance with the flat plate normals associated with target class 2.

A second prominent observation is the low KLD value regions present from upper-left to lower-right in both Figure 27 and Figure 28. These regions indicate transmit-receive angle combinations where the two target classes are similar. The transmit-receive angle combinations in these regions correspond to large bistatic angles. As discussed in Section 3.2, large bistatic angles result in target returns being compressed to just a few range bins. The statistical properties for these returns (i.e. sample mean and sample covariance) provide little distinction between target classes, thereby giving similar KLD values.

A third significant observation is the difference in target classes due to non-dominant scatterers. In both Figure 27 and Figure 28, large diagonal regions of moderate KLD values extend from the upper-left to lower-right of the plot. These regions indicate the potential to distinguish between target classes even if unique, dominant scatterers are not present for a particular target class.

5.3.1.2 PCC vs. KLD

The figures presented in Section 5.3.1.1 suggest specific combinations of platform transmit and receive angles are more likely to result in better classification. This is indeed the case, in general, when a plot of PCC values constructed in the style of the previous section is analyzed. A plot of PCC values computed using (4.10) is shown in Figure 29. Comparing Figure 27 and Figure 29, there appears to be a correspondence between higher KLD and PCC. While the figures do not match exactly, the first three key observations noted in Section 5.3.1.1 are clearly evident in Figure 29. The presence of the target class 1 dominant scatterers, lower KLD values in large bistatic angle regions, and the ability to still distinguish between target classes on some non-

dominant scatterer regions are visible. The complimentary comparison of PCC values for target class 2 versus target class 1, shown in Figure 30, demonstrates the same key behavioral characteristics relative to KLD values.

The relationship between high KLD values and high classification rates is supported further by observing PCC rates for a collection of KLD values within a margin of the peak KLD value. A comparison of PCC and locations of high KLD values are shown in Figure 31. The locations of KLD values above 90%, 75%, and 50% of the peak KLD value are indicated in subplots (b) through (d) of Figure 31, respectively.

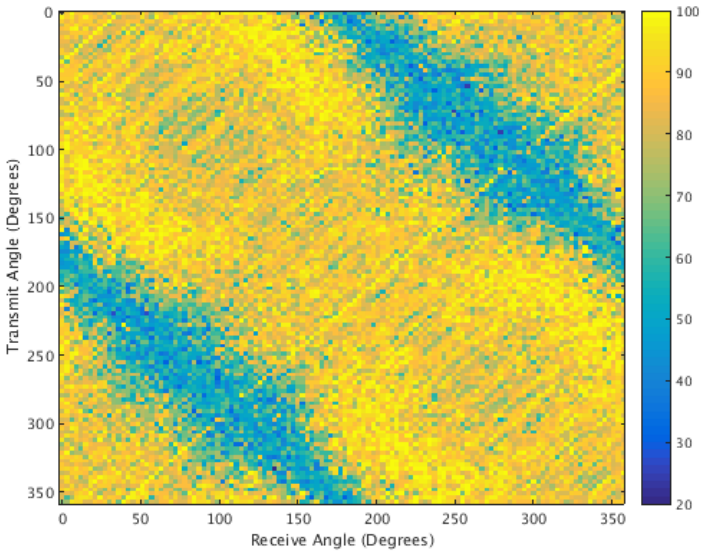


Figure 29. Surface plot showing PCC values comparing target classes 1 and 2, for true target 1, at 9 dB noise power.

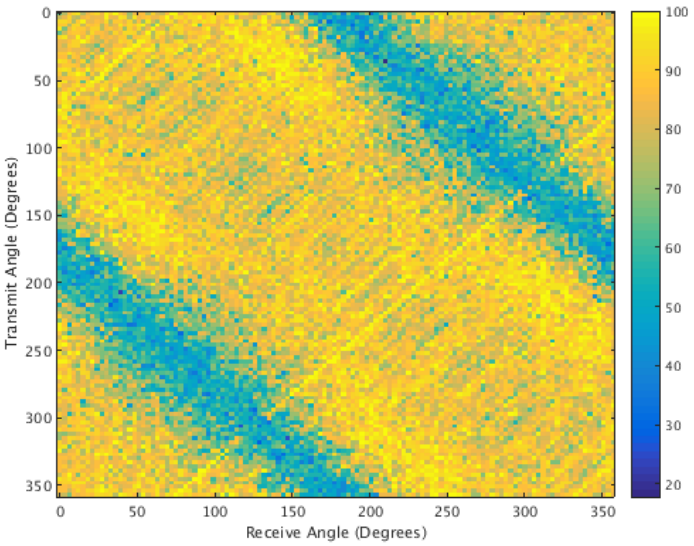
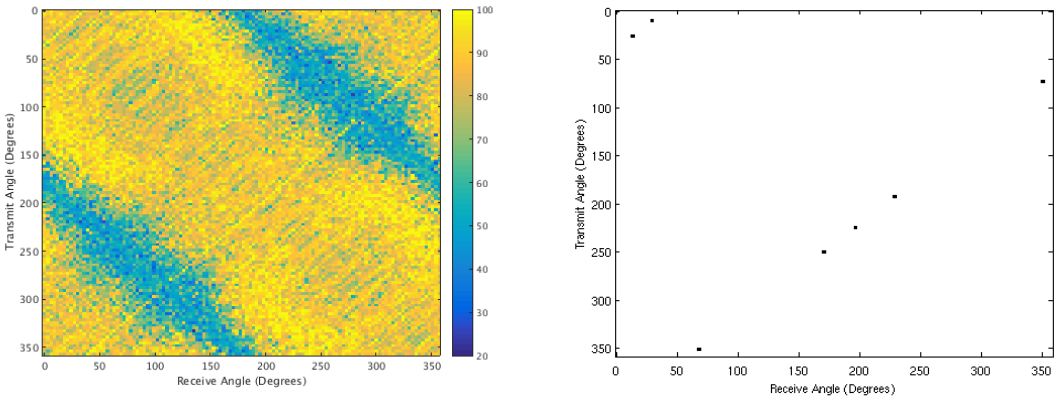
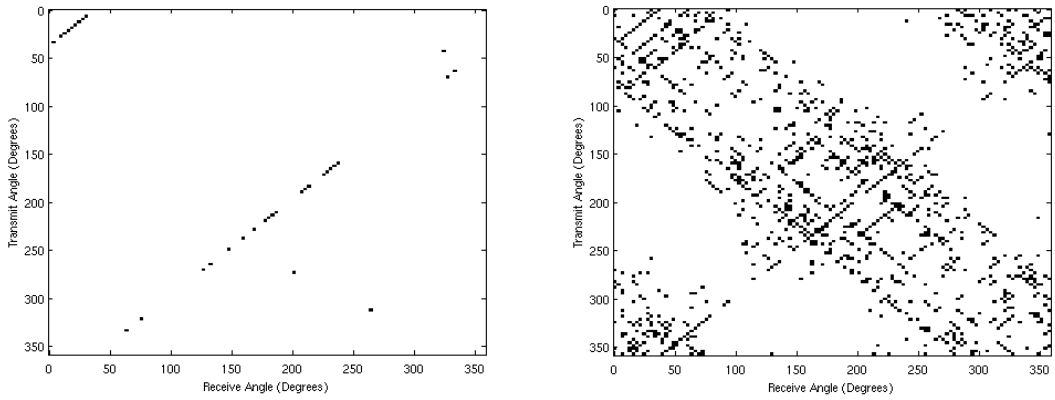


Figure 30. Surface plot showing PCC values comparing target classes 2 and 1, for true target 2, at 9 dB noise power.



(a) **(b)**

Figure 31. Comparison of PCC and high KLD values at different peak thresholds for target class 1 versus target class 2 at 9 dB noise level: (a) PCC, (b) 90% KLD threshold, (c) 75% KLD threshold, and (d) 50% KLD threshold.



(c) **(d)**

Figure 31. Continued.

Two key items are observed in the threshold subplots. First, reflections from the dominant scatterers produce the greatest KLD values, as noted by the upper-right to lower-left diagonals. Second, non-dominant scatterer KLD values show target class separation once the threshold is lowered to 50% of peak. Notably, subplot (d) looks very similar to PCC, as shown in subplot (a). These results might suggest lowering the KLD threshold to 50% would increase the number of transmit-receive angle combinations providing good classification rates. However, this is not the case. An in-depth analysis of the PCC values associated with KLD values identified in subplot (d) indicates a lower probability of having a good classification rate. A histogram of PCC values for transmit-receive angle combinations identified in Figure 31(d) is shown in Figure 32. The histogram reveals good classification rates for a majority of the angle combinations, but less than ninety percent (an arbitrary but reasonable recognition rate) for other angle combinations. The significance of underperforming angle combinations is evident when compared to PCC values for angle combinations derived from dominant scatterer reflectivity only. A histogram of PCC values associated with transmit-receive angle combinations identified in Figure 31 (c) is provided in Figure 33.

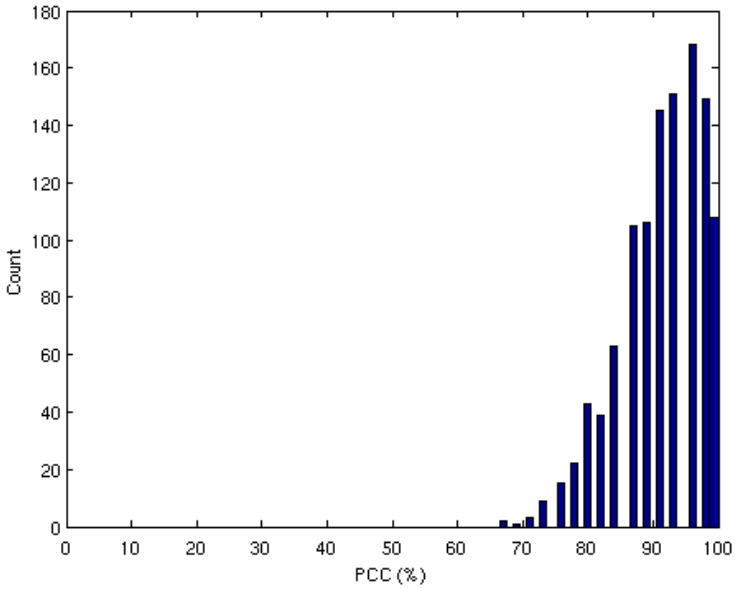


Figure 32. PCC histogram for platform positions associated with target class 1 versus target class 2 KLD values within 50% of peak KLD at 9 dB noise level.

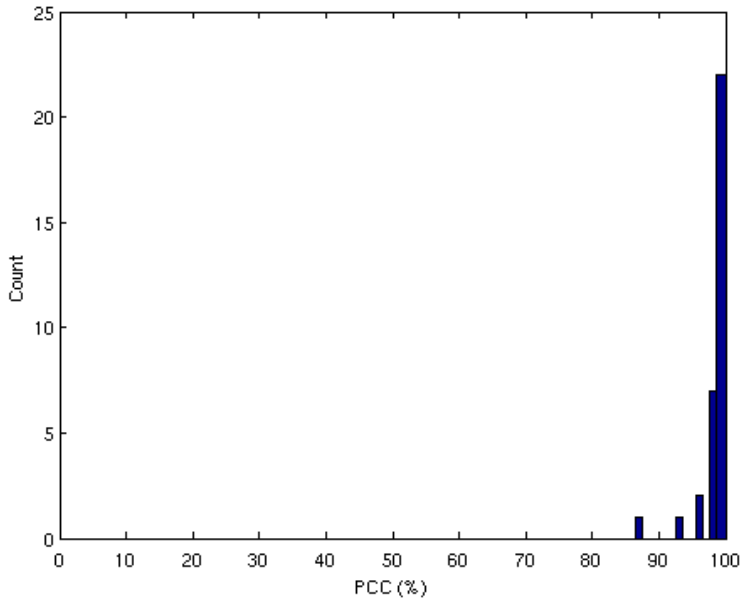


Figure 33. PCC histogram for platform positions associated with target class 1 versus target class 2 KLD values within 75% of peak KLD at 9 dB noise level.

Note the significant improvement in angle combinations producing good classification rates. Specifically, only one out of 33 possible angle combinations produce a PCC value of less than 90%. The comparison of PCC values in Figure 32 and Figure 33 suggest the presence of dominant scatterers can improve classification rates appreciably.

5.3.1.3 Optimal Classification Angle Selection

The results presented in Section 5.3.1.2 illustrate a correspondence between KLD and PCC. This relationship may help identify transmit and receive platform positions that produce the best classification rates. The next step is to demonstrate the effectiveness of KLD in selecting a single, optimal transmit-receive angle combination. In this scenario, (5.3) is used to determine the optimal positions for identifying each target class with one transmit and one receive platform across several noise levels. Two Monte Carlo simulations were executed using the form of (5.3) commensurate with the target class attempting to be identified. The exact form shown in (5.3) is the result of an expectation with respect to target class 1. A complimentary form exists when considering target class 2 by switching i and j. The highest KLD value for target class 1 was achieved for transmit and receive angles of 63 and 333 degrees, respectively. Similarly, the chosen transmit and receive angles for target class 2 were 108 and 324 degrees, respectively.

The corresponding PCC values for these transmit-receive angle combinations are shown in Figure 34, demonstrating excellent performance when compared to classification rates for all possible transmit-receive angle combinations. For example, the target class 1 PCC value for the 27 dB noise level was better than 99.97 % of all other 14,400 possible transmit-receive angle combinations. Similar performance was observed at other noise power levels. A key observation associated with the algorithm's transmit and receive angle selections is that these pairs of angles

are consistent with reflections from the dominant scatterers of each target class. This is the expected result given conclusions drawn in Section 5.3.1.2.

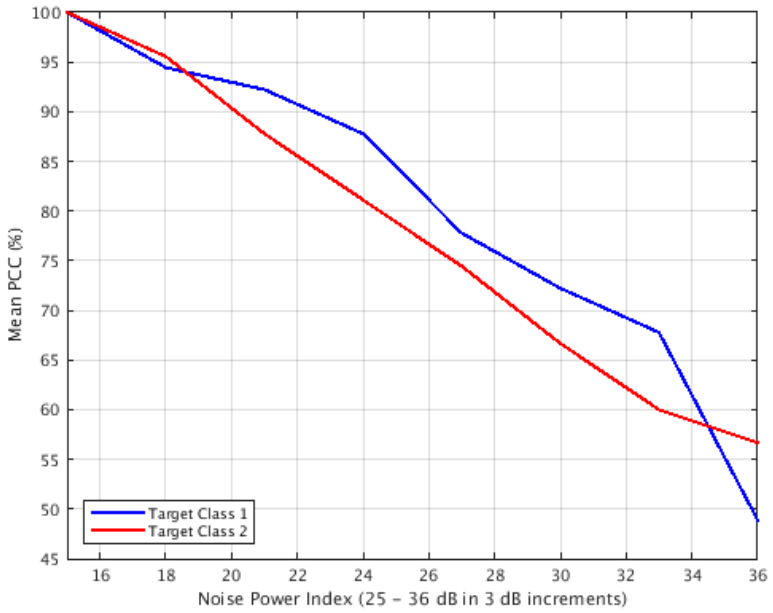


Figure 34. PCC curve for target classes 1 and 2 using transmitter and receiver placements selected by the two-platform KLD.

5.3.2 Single Transmitter, Multiple Receiver Results

An analysis of the optimal and approximate sensor placement results for three platforms were conducted using (5.5) and (5.6). Each sensor placement algorithm was analyzed using a KLD representing different expectation assumptions:

1. KLD values generated with an expectation operation associated with target class 1, KLD_{12};
2. KLD values generated with an expectation operation associated with target class 2, KLD_{21};

3. KLD values associated with a symmetrized KLD calculation referred to as the J-divergence, KLD_{JD}.

The KLD expressions used to identify optimal placements with respect to each individual target class are inconvenient for identifying platform positions where both target classes must achieve a high classification rate. PCC maximization for both target classes requires consideration of the distance between target classes i and j with respect to both expectation assumptions simultaneously. Although it does not enjoy the same theoretical basis as the Kullback-Leibler divergence its terms are drawn from, the J-divergence is a popular choice in conditions requiring this type of symmetrization. As defined in [43], the J-divergence includes KLD expressions from both expectation assumptions:

$$KLD_{JD} = KLD_{ij} + KLD_{ji} = E_i\left[I\left(\hat{\mathbf{z}}\right)\right] + E_j\left[I\left(\hat{\mathbf{z}}\right)\right]. \tag{5.7}$$

Monte Carlo simulations were performed over 13 noise power levels (0 to 36 dB) for each combination of expectation assumption and sensor placement algorithm. Several key observations were apparent when analyzing the suggested placements at each noise level. First, placements at lower noise levels (less than 15 dB) varied considerably with noise level. However, placements became more stable as the noise level increased. This behavior can be attributed to the non-Gaussian nature of the HRR profiles at the lower noise levels. Our KLD expression explicitly assumed the magnitude signals were Gaussian. With less Gaussian noise "corrupting" the HRR profiles at lower noise levels, the probability distributions looked less Gaussian and more Rician. At the higher noise levels, the HRR profile probability distributions looked more Gaussian, which allowed the KLD algorithm to consistently differentiate the targets at the suggested placements. The angles identified in this noise region are presented in Table 8.

Second, the KLD algorithms provided several angle combinations attributable to high classification rates. In these instances, the KLD values for the different angle combinations approached one another asymptotically as the noise level increased. As hypothesized, a detailed analysis of the PCC curves for each of these angle combinations demonstrated equally good classification performance. This observation supports the conclusions drawn in Section 5.3.1.2 regarding the selection of angle combinations within 90% of the peak KLD value.

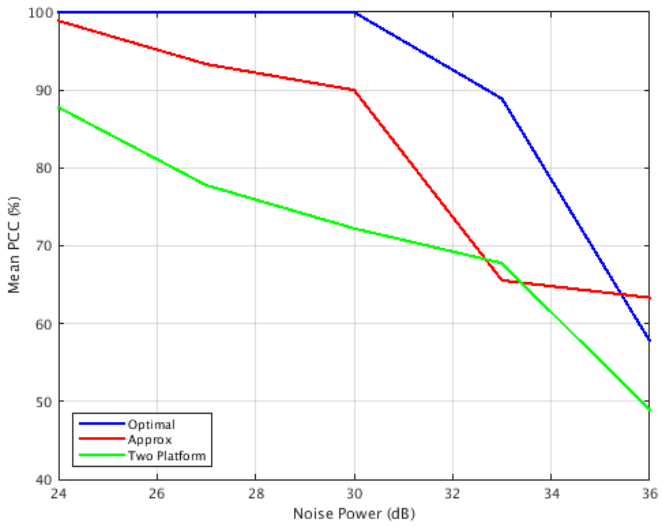
Table 8. Platform transmit and receive angles obtained from the optimal and approximate sensor placement algorithms.

	Optimal			Approximate		
Expectation Assumption	**Transmit Angle (Degrees)**	**Receive Angle 1 (Degrees)**	**Receive Angle 2 (Degrees)**	**Transmit Angle (Degrees)**	**Receive Angle 1 (Degrees)**	**Receive Angle 2 (Degrees)**
Class 1 vs. Class 2	327	69	114	63	333	174
Class 2 vs. Class 1	241	162	171	108	324	72
J-divergence	327	69	114	209	177	318

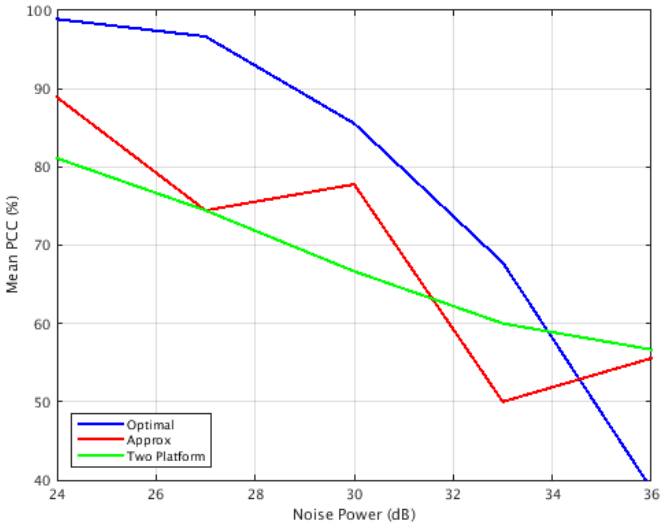
5.3.2.1 Three-platform versus Two-platform Performance

Classifier outputs were computed for each of the placements identified in Table 8. Outputs for the optimal and approximate algorithm placement selections were compared to those produced in Section 5.3.1.3 for the two-platform scenario, as shown in Figure 35. For target class 1, classification rates for the optimal sensor placement exceeded the rates achievable with the two-platform placement by up to 27%. These results lead us to believe, at least in the case of target class 1, that the approximate placement algorithm could be less reliable. This is a reasonable conclusion given the approximate algorithm's restrictions on placing the second receiver. PCC performance for target class 2 is clearly better for the optimal placements when

compared to the equivalent two-platform scenario until PCC rates close to 50% are achieved. We expect inconclusive behavior once rates get close to 50% since the noise level forces the classifier to essentially perform a "coin flip."



(a)

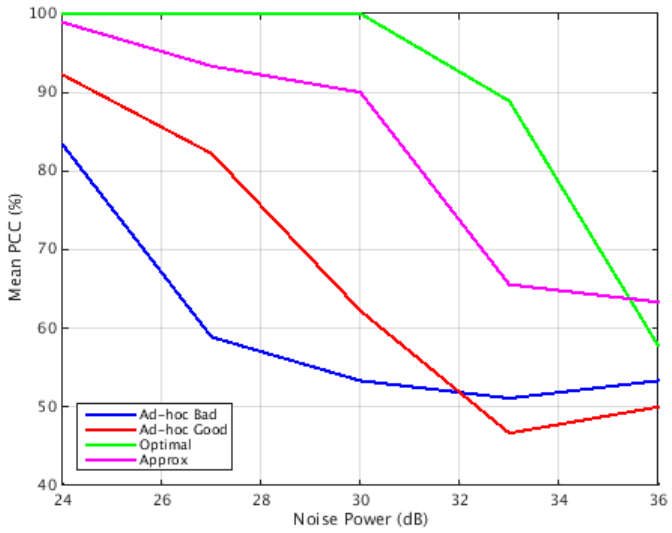


(b)

Figure 35. PCC performance plots showing optimal and approximate sensor placement classification performance for (a) target class 1 and (b) target class 2, relative to two-platform optimal placement performance.

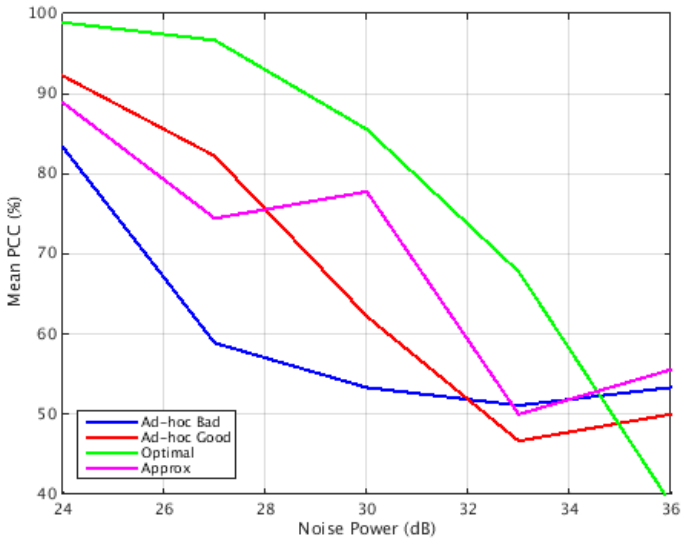
5.3.2.2 *Intelligent versus Ad-hoc Placement Selection*

The classification rates associated with both placement algorithms were also compared to those obtained in Section 4.3 for the ad-hoc sensor placements. This comparison demonstrates improvements attributable to intelligent selection versus ad-hoc sensor placement, as shown in **Error! Reference source not found.**. For both target classes, the optimal and approximate sensor placements exhibited much better classification performance than the ad-hoc placements. Classification rate increases of up to 40% and 20% were achieved for target classes 1 and 2, respectively.



(a)

Figure 36. PCC performance plots showing optimal and approximate sensor placement classification performance, for expectation in the KLD expression taken with respect to (a) target class 1 and (b) target class 2, relative to ad-hoc sensor placement.



(b)

Figure 36. Continued.

These results are even more dramatic when considering the platform configurations employed under both scenarios. In the case of the optimal and approximate placement results, three platforms were used in a single-transmit, dual-receive configuration; however, the ad-hoc placement results used three platforms in a triple-transmit, triple-receive configuration. The optimal and approximate classification results are achieved by using two classifier inputs compared to nine inputs from the ad-hoc placement configuration.

5.3.2.3 Efficacy of the J-divergence Measure

The results presented in Sections 5.3.2.1 and 5.3.2.2 analyzed the performance of the placement selection algorithms in the context of individual target class identification. An analysis of placements yielding good classification rates for both target classes simultaneously required

the J-divergence. Using the J-divergence placement selections shown in Table 8, PCC curves for the optimal and approximate placements were compared to those obtained from the ad-hoc placements. The results shown in Figure 37 demonstrate good performance for the J-divergence placements. Both placement algorithms produce PCC curves exceeding those achieved by ad-hoc placement.

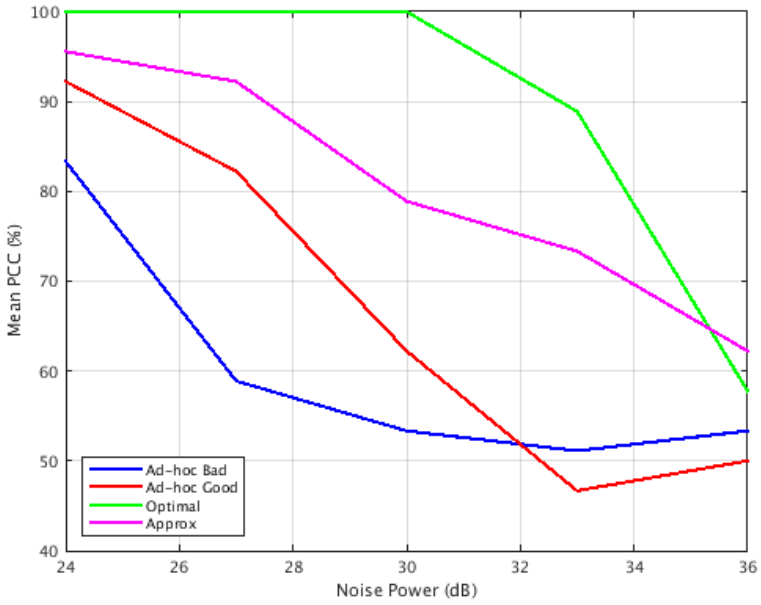


Figure 37. PCC comparison of J-divergence and ad-hoc sensor placements.

5.4 Summary

A distance measure between two likely target classes was used to demonstrate how leveraging statistical differences in target class probability densities can lead to higher classification rates. An approximation of the multivariate Kullback-Leibler distance measure was derived to identify optimal and approximate sensor placements for transmitting and receiving radar platforms. The distance measure was used to demonstrate agreement between high KLD

and PCC values. PCC values associated with the optimal and approximate sensor placement algorithms exceeded up to 99.97% of the remaining transmitter-receiver angle combination PCC values. While transmitter-receiver angle combinations not associated with a dominant scatterer reflectivity path performed reasonably well, angle combinations attributable to dominant scatterer reflectivity performed extremely well.

Several key similarities between KLD and PCC were noted, including the presence of dominant scatterers, lower values in large bistatic angle regions, and the ability to distinguish between target classes in non-dominant scatterer regions. The KLD measure consistently revealed the presence of dominant scatterers, regardless of noise level and expectation operation assumption.

A comparison of the optimal and approximate sensor placement algorithms revealed both algorithms provide excellent classification performance. However, the approximate algorithm demonstrated several performance inconsistencies attributable to the prescribed approximations. Comparison of the single-receive and two-receive platform classification performance results showed the benefits of adding a second receive platform. The superiority of each angle selection algorithm was demonstrated by comparing earlier results for ad-hoc sensor placement to those produced by the intelligent selection algorithm for two receivers. The performance of the intelligent selection algorithm was even more notable given the significantly lower number of classifier inputs required to achieve the observed PCC values. Symmetrization of the KLD for purposes of identifying platform placements yielding good PCC performance for both target classes simultaneously was successful.

6 CONCLUSIONS AND FUTURE WORK

The theoretical results detailed in this dissertation advance the state of the art for ATR. Previous studies in ATR independently demonstrated the benefits of using multiple perspectives and bistatic geometry. The theoretical analysis presented in Chapter 4 merged these two concepts to show how multi-static reflectivity from dominant scatterers can lead to improved target recognition. The inclusion of multiple transmit and receive sensors permitted the exploitation of bistatic geometry. This was leveraged further by deriving an information-theoretic distance measure for determining placement of the transmit and receive sensors. The derived measure exhibited an ability to improve classification rates relative to ad-hoc placement. To our knowledge, there are no previous studies addressing sensor placement for optimizing target classification with multi-static radar.

6.1 Multi-static, Distributed, Multi-sensor Classifier

A detailed explanation and derivation of a multi-static, distributed, multi-sensor classifier was presented in Chapter 4. The classifier was subjected to several critical analyses to determine its accuracy and robustness. A comparative analysis was performed to demonstrate that increases in target classification rates were attributable to bistatic geometry and not due to the mere addition of sensors. For the scenarios considered, the simulations revealed PCC improvements of approximately 25% were achievable by adding additional monostatic sensors. However, the approximate algorithm generally did not perform as well as the optimal algorithm.

Contributions of dominant scatterers were shown to improve classification rates by as much as 30% over similar tests with targets lacking dominant scatterers. These results were expected and formed the basis for the exemplary findings associated with the sensor placement

algorithms. A series of simulations was performed to assess the classifiers as a function of mismatches in target pose angle assumptions. Classification performance curves showed expected behavior as the uncertainty in target pose angle transitioned from deterministic to Gaussian to uniform. In cases where target pose angle uncertainty was assumed to be Gaussian, matching the classifier's assumed pose angle error (i.e., standard deviation) to the actual pose angle error provided the best classification results.

Classifier tests were also executed to analyze performance related to modified target modeling camouflage, concealment, and deception. Monostatic receiver classification rates illustrated the difficulty in successfully identifying modified targets. The multi-sensor classifier demonstrated superior performance in identifying the modification. Classification rates within 10% of the unmodified target classification rates were achieved by the multi-sensor classifier.

An assessment of the classifier's performance related to different transmitter and receiver configurations was performed. Specifically, two multi-sensor configurations were tested to assess classification performance for single transmitter, stand-off and multiple-transmit, multiple-receive scenarios. Classification performance for the single transmitter scenario, while not as good as the multiple-transmit scenario, showed promise in providing good classification. Both configurations exhibited excellent performance at lower SNR regions with improvements up to 30% over monostatic configurations.

Finally, a simple test of the classifier's performance for two ad-hoc sensor placements was completed. This test revealed large variances in classifier performance based upon sensor placement, demonstrating the necessity of identifying a suitable sensor placement algorithm.

6.2 Sensor Placement for Improved Target Classification

Works by Kullback and Leibler [43] provided the foundation for developing sensor placement algorithms addressing multi-sensor classification. An information-theoretic approach was taken to derive two sensor placement algorithms. These algorithms used a multi-sensor approximation of the Kullback-Leibler distance measure to select sensor positions producing high classification rates. The distance measure worked well in a simple single-transmit, single-receive scenario. The sensor placements suggested in this simple scenario consistently produced high classification rates. An example for the 24 dB noise level was provided wherein the platform placements produced a PCC greater than 99.97% of all other possible platform placement combinations.

Both sensor placement algorithms demonstrated good utility in several three-platform tests. In these tests, a single transmitter and two receivers were subjected to both placement prediction algorithms. Suggested placements from the optimal algorithm resulted in individual target class PCC rates exceeding similar tests in the two-platform scenario. Suggested placements from the approximate algorithm produced less conclusive, but good, PCC rates. These tests validated the benefits of having a second receive platform and the usefulness of the placement selection algorithms.

A comparison of the PCC curves for each placement algorithm to those obtained for ad-hoc sensor placement was also performed. Despite having one less receive sensor and using only a single transmitter, the single-transmit, dual-receive configuration with intelligent placement outperformed the ad-hoc placements. Attempts to maximize PCC values for both target classes simultaneously were also successful; noting PCC curves from both algorithms' platform placements exceeded those from the ad-hoc placements.

6.3 Future work

The work presented in this dissertation provides a basis for further study of multi-static, distributed, multi-sensor ATR. The classifier and distance measure results presented in Chapters 4 and 5 were derived from a two-dimensional geometric model. Extension of the geometric model to three dimensions, possibly employing a robust EM code such as XPATCH or Lucernhammer, would offer a more complete analysis. An analysis of the impact of limited training data would also prove useful. The simulations included in this thesis assumed training data sets for all transmit and receive angles were available to the classifier. It is reasonable to assume that a limited training data set would influence PCC and KLD. Another significant enhancement to this effort would be the consideration of more than two probable target classes in the KLD analysis. Other potential topics include the effects of the HRR profile range window on classification performance and a Cramer-Rao bound analysis on pose estimation. Finally, potential application of the derived algorithms to other target recognition venues would be interesting. For example, are the algorithms suitable for classifying human and airborne objects?

A. APPENDIX – DERIVATION OF A MULTI-SENSOR KULLBACK-LEIBLER DISTANCE MEASURE

Detailed below is the derivation of the Kullback-Leibler distance measure between target classes 1 and 2 calculated with respect to the target class 1 probability distribution. The final expressions are actually a lower bound on the *approximate* Kullback-Leibler distance, but for brevity, we omit the modifier "a lower bound on the approximate" through the text of this dissertation and simply refer to it as "the KLD". A similar derivation can be performed to produce a lower bound on the approximate Kullback-Leibler distance measure between the same target classes but with respect to the target class 2 probability distribution.

From (5.2),

$$\begin{aligned} E_i\left[I(\hat{\mathbf{z}})\right] &= E_{p\left(\hat{\mathbf{z}}|H_i;\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)}\left[\ln\left(\frac{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_i;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1},\ldots,\hat{\mathbf{z}}_{A,B}\mid H_j;\boldsymbol{\theta}_{1,1},\cdots,\boldsymbol{\theta}_{A,B}\right)}\right)\right] \\ &= E_{p\left(\hat{\mathbf{z}}|H_i;\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)}\left[\ln\left(\frac{p\left(\hat{\mathbf{z}}_{1,1}\mid H_i;\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_i;\boldsymbol{\theta}_{A,B}\right)}{p\left(\hat{\mathbf{z}}_{1,1}\mid H_j;\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_j,\boldsymbol{\theta}_{A,B}\right)}\right)\right] \end{aligned}. \quad (A.1)$$

Assuming the prior densities for target pose angle $\theta_{pose}^{'}$ and $\theta_{pose}^{''}$ are uniform, and assuming that the integrals can be reasonably approximated by their maximum value:

$$E_{p\left(\hat{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\ln\left(\frac{\max\limits_{\theta_{pose}^{'}}\left\{\exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{A,B}\right)\right]\right\}\right\}}{\max\limits_{\theta_{pose}^{''}}\left\{\exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_j,\theta_{pose}^{''};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_j,\theta_{pose}^{''};\boldsymbol{\theta}_{A,B}\right)\right]\right\}\right\}}\right)\right]. \quad (A.2)$$

Here, replacing an integral with its maximum value can be thought of as approximating the density by a single Dirac impulse function sampling the density at its maximum point. This is a rough adaptation of the approach to nuisance parameters taken by Grenander, Srivastava, and

Miller [46], who used Laplace's method to integrate posterior densities over nuisance pose parameters by approximating them as Gaussian;[6] we essentially further approximate this Gaussian as being an impulse. The quantity inside the expectation in (A.2) is similar in spirit to the first term in Equation (8) of [46]. The authors of [46] have an additional term involving the Fisher information on the pose parameters that provides a more direct link between classifier performance and the ability to estimate nuisance parameters; incorporating such a term into the derivation provided here, and exploring what effects it would have, is left as a direction for future research.

We further approximate the expression by assuming that the maximizing θ' in the numerator is θ_{pose}^{true}. This is analogous to the application of Proposition 1 on p. 1661 to arguments in the left column of p. 1662 of [46] and the "asymptotic situation" result in Section III.A of [46]. The maximization of θ_{pose}'' in the denominator can be intuitively thought of as the pose that makes the confusing target "look most like" the true target; a similar argument is made in Section V of [46]. In our case, we use:

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\ln\left(\frac{\exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{A,B}\right)\right]\right\}}{\max\limits_{\theta_{pose}}\left\{\exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_j,\theta_{pose}'';\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_j,\theta_{pose}'';\boldsymbol{\theta}_{A,B}\right)\right]\right\}\right\}}\right)\right] \quad \text{(A.3)}$$

Since the exponential function is monotonic, we may interchange the order of operations between the maximum and exponential functions in the denominator:

[6] Their use of a Gaussian approximation on full posterior density for the nuisance pose parameter should not be confused with the Gaussian likelihood of the data itself; even if the probability of the data conditioned on a particular pose is truly Gaussian, the mixture of Gaussians resulting from integrating out pose is not.

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\ln\left(\frac{\exp\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{A,B}\right)\right]\right\}}{\exp\left\{\max_{\theta_{pose}^{"}}\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_j,\theta_{pose}^{"};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_j,\theta_{pose}^{"};\boldsymbol{\theta}_{A,B}\right)\right]\right\}\right\}}\right)\right]. \quad \text{(A.4)}$$

Applying the natural logarithm function to both numerator and denominator:

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\begin{array}{l}\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_i,\theta_{pose}^{true};\boldsymbol{\theta}_{A,B}\right)\right]\\ -\max_{\theta_{pose}^{"}}\left\{\ln\left[p\left(\hat{\mathbf{z}}_{1,1}\mid H_j,\theta_{pose}^{"};\boldsymbol{\theta}_{1,1}\right)\cdots p\left(\hat{\mathbf{z}}_{A,B}\mid H_j,\theta_{pose}^{"};\boldsymbol{\theta}_{A,B}\right)\right]\right\}\end{array}\right]. \quad \text{(A.5)}$$

Substituting Gaussian multivariate probability density expressions into (A.5) and discarding constants:

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\begin{array}{l}\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left[\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right]\right.\right.\\ \quad\left.\left.-\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right\}\right\}\\ \quad-\max_{\theta_{pose}^{"}}\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left[\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right]\right.\\ \quad\left.-\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right\}\end{array}\right]. \quad \text{(A.6)}$$

The dependence of $\hat{\mathbf{z}}_{a,b}$ on $\boldsymbol{\theta}_{a,b}$ and θ_{pose}^{true} has been suppressed for clarity. Recognizing the quadratic term as a scalar and using the property that the trace of a scalar is equal to the scalar:

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\begin{array}{l}\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left[\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right]\right.\right.\\ \quad\left.\left.-tr\left\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right\}\right\}\right\}\\ \quad-\max_{\theta_{pose}^{"}}\left\{\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left[\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right]\right.\right.\\ \quad\left.\left.-tr\left\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right\}\right\}\right\}\end{array}\right]. \quad \text{(A.7)}$$

Using the property of cycling matrices within the trace:

$$E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\Bigg[\sum_{a=1}^{A}\sum_{b=1}^{B}\Big\{-\ln\Big[\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\Big] - tr\Big\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\Big\}\Big\} - \max_{\theta_{pose}^{"}}\Big\{\sum_{a=1}^{A}\sum_{b=1}^{B}\Big\{-\ln\Big[\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\Big] - tr\Big\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\Big\}\Big\}\Big\}\Bigg] \quad \text{(A.8)}$$

Since the expectation function is a linear operation, we can bring the expectation operator inside the summations in the first term of (A.8). To treat the second term the same way requires exchanging the order of the maximum and expectation operations; in general, this cannot be done while retaining equality, but we can exchange the operations and form a lower bound:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\Big[-\ln\Big(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\Big) - tr\Big\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\Big\}\Big] - \max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\Big[-\ln\Big(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\Big) - tr\Big\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\Big\}\Big] \quad \text{(A.9)}$$

Expression (A.9) is a lower bound on the approximate KLD in (A.8). If the likelihood function is fairly peaked, as we have often observed, we conjecture that this bound is fairly tight. Further study of the tightness of this bound, as well as the accuracy of our KLD approximations in general, is left as an avenue for future study.

Recognizing that the determinant terms are constants with respect to the expectation operation:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}\Bigg\{-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right) - E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[tr\left\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right\}\right]\Bigg\} - \max_{\theta_{pose}^{"}}\Bigg\{\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right) - E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[tr\left\{\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right\}\right]\Bigg\} \quad . \quad (A.10)$$

Since the trace operator is linear, we can swap the expectation and trace operations:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}\Bigg\{-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right) - tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right]\right\}\Bigg\} - \max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\Bigg\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right) - tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right]\right\}\Bigg\} \quad . \quad (A.11)$$

The covariance term $\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)$ is constant with respect to the expectation, and the expectation of the quadratic term $\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_k\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)^T$ is just the covariance matrix of $\hat{\mathbf{z}}_{a,b}$, yielding

$$\sum_{a=1}^{A}\sum_{b=1}^{B}\Bigg\{-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right) - tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right]\right\}\Bigg\} - \max_{\theta_{pose}^{"}}\Bigg\{\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right) - tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right]\right\}\Bigg\} \quad . \quad (A.12)$$

The matrix product $\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\mathbf{C}_i^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)$ is just the identity matrix. Applying the expectation and trace operator to this quantity results in P, the dimension of $\hat{\mathbf{z}}_{a,b}$:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)-P$$
$$-\max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)\right. \quad \text{. (A.13)}$$
$$\left.-tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\left(\hat{\mathbf{z}}_{a,b}-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right]\right\}\right\}$$

Expanding the expectation operation on the remaining quadratic term:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)-P-\max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)\right.$$
$$-tr\left\{E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\hat{\mathbf{z}}_{a,b}\hat{\mathbf{z}}_{a,b}^T-\hat{\mathbf{z}}_{a,b}\boldsymbol{\mu}_j^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\hat{\mathbf{z}}_{a,b}^T+\right.\right. \quad \text{(A.14)}$$
$$\left.\left.\left.+\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\boldsymbol{\mu}_j^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right]\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right\}\right\}$$

Distributing the expectation operation to all expanded quadratic terms:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)-P-\max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)\right.$$
$$-tr\left\{\left[R_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\boldsymbol{\mu}_j^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right.\right. \quad \text{, (A.15)}$$
$$\left.\left.\left.-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\boldsymbol{\mu}_i^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)+\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\boldsymbol{\mu}_j^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right]\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right\}\right\}$$

where $R_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)$ is the correlation matrix for $\hat{\mathbf{z}}_{a,b}$, $E_{p\left(\tilde{\mathbf{z}};\boldsymbol{\theta}_{a,b},\theta_{pose}^{true},H_i\right)}\left[\hat{\mathbf{z}}_{a,b}\hat{\mathbf{z}}_{a,b}^T\right]$. Adding and subtracting $\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\boldsymbol{\mu}_i^T\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)$ to the terms inside the trace allows simplification to a more relevant formulation of the trace term. The correlation term $R_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)$ becomes $C_i\left(\boldsymbol{\theta}_{m,n},\theta_{pose}^{true}\right)$ and the quadratic mean terms can be simplified to a more compact quadratic form:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)-P-\max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)\right.$$
$$-tr\left\{\left[\mathbf{C}_{\mathbf{i}}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)+\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right. \quad . \quad \text{(A.16)}$$
$$\left.\left.\cdot\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\right]\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right\}\right\}$$

Reintroducing $\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)$ into the trace term gives a simplified final form:

$$\sum_{a=1}^{A}\sum_{b=1}^{B}-\ln\left(\det\left(\mathbf{C}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\right)\right)-P-\max_{\theta_{pose}^{"}}\sum_{a=1}^{A}\sum_{b=1}^{B}\left\{-\ln\left(\det\left(\mathbf{C}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right)\right.$$
$$-tr\left\{\mathbf{C}_{\mathbf{i}}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right\} \quad . \quad \text{(A.17)}$$
$$\left.+\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)^T\cdot\mathbf{C}_j^{-1}\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\left(\boldsymbol{\mu}_i\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{true}\right)-\boldsymbol{\mu}_j\left(\boldsymbol{\theta}_{a,b},\theta_{pose}^{"}\right)\right)\right\}$$

REFERENCES

[1] Shaw, A. K.; Paul, A. S.; Williams, R., "Eigen-templated-based HRR-ATR with Multi-look and Time-recursion," *IEEE Trans. AES,* vol. 49, no. 4, pp. 2369-2385, 2013.

[2] Paul, A. S.; Shaw, A. K., "Robust HRR Radar Target Identification by Hybridization of HMM and Eigen-template-based Matched Filtering," in *Proc. SPIE,* vol. 5094, pp. 278-289, 2003.

[3] Srinivas, U.;Monga, V.;Riasati, V., "A Comparative Study of Basis Selection Techniques for Automatic Target Recognition," in *Proc. 2012 IEEE Radar,* pp. 710-713, 2012.

[4] Bhatnagar, V.;Shaw, A. K.;Williams, R., "Improved Automatic Target Recognition Using Singular Value Decomposition," in *Proc. 1998 IEEE ASSP,* pp. 2717-2720, 1998.

[5] Nelson, D. E.;Starzyk, J. A., "High Range Resolution Radar Signal Classification - A Partitioned Rough Set Approach," in *Proc. 2001 IEEE System Theory,* pp. 21-24, 2001.

[6] Mingjing, L.;Zhefeng, Z.;Ming, H., "Radar Target Recognition Based on Combined Features of High Range Resolution Profiles," in *Proc. 2009 ASPAR Synthetic Aperture Radar,* pp. 876-879, 2009.

[7] Nelson, D. E.;Starzyk, J. A., "Advanced Feature Selection Methodology for Automatic Target Recognition," in *Proc. 1997 System Theory Sym.,* pp. 24-28, 1997.

[8] H. V. Poor, An Introduction to Signal Detection and Estimation, New York: Springer, 1988.

[9] S. Wong, "Multi-look Fusion Identification: A Paradigm Shift From Quality to Quantity in Data Samples," in *Proc. SPIE Automatic Target Recognition XIX,* vol. 7335, pp. G1-G12, 2009.

[10] Nguyen, D. H.;Kay, J. H.;Orchard, B. J.;Benitz, G. R.;Whiting, R. H., "Improving HRR ATR Performance at Low SNR by Multi-look Adaptive Weighting," in *Proc. SPIE Automatic Target Recognition XI,* vol. 4379, pp. 216-228, 2001.

[11] Stinco, P.;Greco, M.;Gini, F.;La Manna, M., "Multistatic Target Recognition in Real Operational Scenarios," in *Proc. 2012 IEEE Radar Conf.,* pp. 354-359, 2012.

[12] Jacobs, S. P.;O'Sullivan, J. A., "Automatic Target Recognition Using Sequences of High Resolution Radar Range Profiles," *IEEE Trans. AES,* vol. 36, no. 2, pp. 364-382, 2000.

[13] Ma, J.;Lie, H.;Ahalt, S. C., "Using Support Vector Machines as HRR Signature Classifiers," in *Proc. SPIE Automatic Target Recognition XI,* vol. 4379, pp. 209-215, 2001.

[14] Vespe, M.;Baker, C. J.;Griffiths, H. D., "Perspective Aperture Dependency of Netted Radar Target Classification," in *Proc. EMRS DTC 3rd Annual Technical Conference,* vol. 18, no. 2, 2006.

[15] Vespe, M.;Baker, C. J.;Griffiths, H. D., "Radar Target Classification Using Multiple Perspectives," *IET Radar, Sonar, and Navigation,* vol. 1, no. 4, pp. 300-307, 2007.

[16] Rizvi, S. A.;Nasrabadi, N. M., "Fusion Techniques for Automatic Target Recognition," in

Proc. IEEE 32nd Applied Imagery Pattern Recognition Workshop, pp. 27-32, 2003.

[17] Kahler, B.;Blasch, E., "Robust Multi-look HRR ATR Investigation Through Decision-level Fusion Evaluation," in *Proc. 11th Information Fusion Conference,* pp. 1-8, 2008.

[18] Zhang, H.;Nasrabadi, N. M.;Zhang, Y.;Huang, T. S., "Multi-view Automatic Target Recognition Using Joint Sparse Representation," *IEEE Trans. AES,* vol. 48, no. 3, pp. 2481-2495, 2012.

[19] Vespe, M.;Baker, C. J.;Griffiths, H. D., "Automatic Target Recognition Using Multi-diversity Radar," *IET Radar, Sonar, and Navigation,* vol. 1, no. 6, pp. 470-478, 2007.

[20] Cui, J.;Gudnason, J.;Brookes, M., "Hidden Markov Models for Multi-perspective Radar Target Recognition," in *Proc. 2008 IEEE Radar Conf.,* pp. 1-5, 2008.

[21] Butler, T. B.;Goodman, N. A., "Multistatic Target Classification with Adaptive Waveforms," in *Proc. 2008 IEEE Radar Conf.,* pp. 1-6, 2008.

[22] M. A. Richards, "Elements of a Pulsed Radar," in *Fundamentals of Radar Signal Processing*, New York, McGraw-Hill, 2005, p. 6.

[23] H. L. V. Trees, "Range Spread Targets and Channels," in *Detection, Estimation, and Modulation Theory*, Malabar, Krieger Publishing Company, 1992, pp. 413-443.

[24] H. L. V. Trees, "Range-spread Model," in *Detection, Estimation, and Modulation Theory*, Malabar, Krieger Publishing Company, 1971, p. 414.

[25] M. A. Richards, "Statistical Description of Radar Cross Section," in *Fundamentals of Radar*

Signal Processing, New York, Springer, 2005, pp. 67-79.

[26] Knott, E. F.;Shaeffer, J. F.;Tuley, M. T., "IEEE RCS Definition," in *Radar Cross Section*, Edison, SciTech, 2004, p. 64.

[27] Naval Air Systems Command, "Radar Cross Section," in *Electronic Warfare and Radar Systems Engineering Handbook*, Washington D.C., Naval Air Systems Command, 1999.

[28] Knott, E. F.;Shaeffer, J. F.;Tuley, M. T., "High Frequency RCS Prediction Techniques," in *Radar Cross Section*, Edison, SciTech, 2004, pp. 183-224.

[29] Jackson, J. A.;Rigling, B. D.;Moses, R. L., "Parametric Scattering Models for Bistatic Synthetic Aperture Radar," *in Proc. 2008 IEEE Radar Conf.,* pp. 1-5, 2008.

[30] S. M. Kay, "Summary of Composite Hypothesis Testing," in *Fundamentals of Statistical Signal Processing Volume II*, Upper Saddle River, Prentice Hall, 1998, p. 187.

[31] Fishler, E.;Haimovich, A.;Blum, R. S.;Cimini, L. J.;Chizhik, D.;Valenzuela, R. A., "Spatial Diversity in Radars - Models and Detection Performance," *IEEE Trans. Signal Processing,* vol. 54, no. 3, pp. 823-838, 2006.

[32] Godrich, H.;Haimovich, A. M.;Blum, R. S., MIMO Radar Signal Processing, Hoboken: John Wiley & Sons, 2009.

[33] Bae, J.;Goodman, N. A., "Widely Separated MIMO Radar with Adaptive Waveform for Target Classification," in *Proc. 4th IEEE International Workshop on CAMSAP,* pp. 21-24, 2011.

[34] H. L. V. Trees, "Model and Intuitive Discussion," in *Detection, Estimation, and Modulation Theory*, Malabar, Krieger Publishing Company, 1971, p. 415.

[35] Rigling, B.;Moses, R. L., "GTD-based Scattering Models for Bistatic SAR," in *Proc. SPIE Algorithms for Synthetic Aperture Radar Imagery XI,* vol. 5427, pp. 208-219, 2004.

[36] Beaulieu, N. C.;Hemachandra, K. T., "Novel Simple Forms for Multivariate Rayleigh and Rician Distributions with Generalized Correlation," in *Proc. 2010 IEEE Globecom,* pp. 1-6, 2010.

[37] S. M. Kay, "Composite Hypothesis Testing," in *Fundamentals of Statistical Signal Processing - Detection Theory*, Upper Saddle River, Prentice Hall, 1998, pp. 191-197.

[38] M. A. Richards, "Correlation Length," in *Fundamentals of Radar Signal Processing*, New York, McGraw-Hill, 2005, p. 78.

[39] Merhav, N.;Guo, D.;Shamai, S., "Statistical Physics of Signal Estimation in Gaussian Noise: theory and Examples of Phase Transitions," vol. 56, no. 3, pp. 1400-1416, 2010.

[40] M. A. Richards, "Data Integration," in *Fundamentals of Radar Signal Processing*, New York, McGraw-Hill, 2005, pp. 33-35.

[41] Brendel, G. F.;Horowitz, L. L., "Benefits of Aspect Diversity for SAR ATR: Fundamental and Experimental Results," in *Proc. SPIE Algorithms for SAR Imagery VII,* vol. 4053, pp. 567-578, 2000.

[42] T. Kailath, "The Divergence and Bhattacharyya Distance," *IEEE Trans. Communication*

Technology, vol. 15, no. 1, pp. 52-60, 1967.

[43] Kullback, S.;Leibler, R. A., "On Information and Sufficiency," *The Annals of Mathematical Statistics,* vol. 22, no. 1, pp. 79-86, 1951.

[44] Cover, T. M.;Thomas, J. A., "Entropy, Relative Entropy, and Mutual Information," in *Elements of Information Theory*, Hoboken, John Wiley & Sons, 2006, pp. 13-55.

[45] C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal,* vol. 27, pp. 379–423, 623–656, 1948.

[46] Grenander, U.; Srivastava, A.; Miller, M. I., "Asymptotic Performance Analysis of Bayesian Target Recognition," *IEEE Trans. on Information Theory,* vol. 46, no. 4, pp. 1658-1665, 2000.

[47] M. A. Richards, "The Matched Filter," in *Fundamentals of Radar Signal Processing*, New York, Springer, 2005, pp. 161-163.

[48] C. L. Eisenbies, "Classification of Ultra High Range Resolution Radar Using Decision Boundary Analysis," Air Force Institute of Technology, Wright-Patterson AFB, 1994.

[49] S. M. Kay, "Waveform Design for Multistatic Radar Detection," *IEEE Trans. AES,* vol. 45, no. 3, pp. 1153-1166, 2009.

[50] M. A. Richards, "The Stepped Frequency Waveform," in *Fundamentals of Radar Signal Processing*, New York, McGraw Hill, 2005, pp. 206-211.

[51] DeVore, M. D.; Lanterman, A. D.; O'Sullivan, J. A., "ATR Performance of a Rician Model

for SAR Images," *SPIE,* vol. 4050, pp. 34-45, 2000.

VITA

John S. Wilcher received a B.S. in Electrical Engineering from the University of Maryland in May of 1988. While studying at the University of Maryland, he was employed by the United States Department of Navy as an electrical engineering co-operative education student. He performed various engineering duties related to submarine sonar and electrical systems during his time at the Department of Navy. Upon graduation from the University of Maryland, Mr. Wilcher joined the Federal Systems Division of International Business Machines (IBM) as an Associate Systems Engineer in the sonar trainer department. During his tenure at IBM, he worked on several high profile systems related to submarine sonar and surveillance. In addition to be being employed full-time at IBM, he completed his Master in Business Administration (MBA) degree at George Mason University in 1994. Mr. Wilcher left IBM in 1994 to pursue software systems development opportunities in commercial business. For the next several years, he worked for large enterprises including Deloitte & Touche, MCI Telecommunications, Zurich, and Borland. After managing numerous teams on dozens of enterprise-wide software system projects, he entered the small business industry. He participated in several start-up companies, helping instill sound information technology practices and principles into each organizations' business operations. Having gained valuable experience in the business start-up arena, Mr. Wilcher started his own software development company in 2005. The company remains profitable today by providing custom, commercial, back-office software systems for clients in various industries. Once the company was established, he realized his true passion was solving engineering problems within the defense industry. In 2010, he returned to college as a PhD student at the Georgia Institute of Technology. He received his M.S. in Electrical and Computer Engineering in May 2012. Mr. Wilcher has been employed as a graduate research assistant with the Georgia Tech Research Institute since 2010. His research efforts have focused upon radar systems, specifically automatic target recognition. He expects to graduate with his PhD degree in May 2015.