



THESIS

Coherence-based Underwater Target Detection for Side-Scan Sonar Imagery

Submitted by

James Derek Tucker

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2009

COLORADO STATE UNIVERSITY

February 27, 2009

WE HEREBY RECOMMEND THAT THE **THESIS** PREPARED UNDER OUR SUPERVISION BY **JAMES DEREK TUCKER** ENTITLED **COHERENCE-BASED UNDERWATER TARGET DETECTION FOR SIDE-SCAN SONAR IMAGERY** BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

---

Prof. Rockey J. Luo

---

Prof. F. Jay Breidt

---

Prof. Mahmood R. Azimi-Sadjadi  
**Adviser**

---

Prof. Anthony A. Maciejewski  
**Department Head**

## ABSTRACT OF THESIS

### Coherence-based Underwater Target Detection for Side-Scan Sonar Imagery

Detection and classification of underwater objects in sonar imagery is a complicated problem due to various factors such as variations in the operating and environmental conditions, presence of spatially varying clutter, variations in target shapes, compositions, and orientation. Also contributing to the difficulty of the problem is the lack of *a priori* knowledge about the shape and geometry of new non-mine-like objects that may be encountered, as well as changes in the environmental or operating conditions encountered during data collection. The use of *coherent* information is proposed as an excellent tool for detection and classification due to the fact when a target is present in a region of interest within an image not only coherence changes allowing for detecting the target, but also the way it changes is different than a non-target hence enabling the target's classification.

The problem of detection of underwater objects in sonar imagery can be categorized into two types. In the first case, only one sensor is used while the second case involves dual disparate sensors both observing the same event but at perhaps disparate locations, frequency, resolution, etc. In both cases coherent-based detection can be exploited though in different forms.

To implement the coherent-based detector in these cases new formulations are needed to extend the standard Gauss-Gauss detection to the two channel case that can be used for both categories of target detection problems using sonar imagery. This extension requires finding new expressions for the log-likelihood function and J-divergence when the covariance matrices under the different hypothesis are different. In this thesis, we have addressed this problem by developing a new two-channel Gauss-Gauss detection framework that can be applied to single sonar as well as dual

sonar platforms. This is done by relating the two-channel detector to the one-channel case (standard). When applied to single sonar platforms channels are consecutive columns in a region of interest and coherence between the columns provides indication of the presence of a target. For dual disparate sonars, the channels are blocks from a pair of co-registered region of interests and coherence between the pair of blocks provides indication of the presence of a target. The use of dual sonar types, e.g. a high frequency high resolution sonar with good target definition and a low resolution broadband sonar with good clutter suppression ability significantly improves the detection and false alarm rates over the single sonar case. A novel distributed detection system that exploits multiple dual platforms detectors for more than two sonar platforms situations is also developed. Distributed detection system provides much better probability of detection and false alarm rates comparing to single or dual platform cases, due to the fusion of multiple decisions being made about the observations from the environment.

A comprehensive study is carried on the two-channel coherence-based detector and a comparison is made on the detection and false alarm rate performance for three different implementation cases on two data sets acquired from the Naval Surface Warfare Center (NSWC) in Panama City, FL. The first data set contains high frequency side scan sonar images obtained from one sonar with varying degree of difficulty and bottom clutter. The second data set contains multiple sonar imagery, namely one high frequency sonar and three broadband sonars registered over the same region of the target field with varying degree of bottom clutter. Results illustrating the effectiveness of the proposed detection tools are presented in terms of correct detection and false alarm rates for various bottom difficulty conditions. It is observed that the distributed detection that uses two dual disparate implementations provided the best overall detection performance with 53 out of 53 detections and only 7 false alarms per image.

In realistic situations, only a limited number of data samples can be drawn from the environment to design the Gauss-Gauss detector. To determine the effects of the detector's performance in such cases new expressions for the log-likelihood and "signal-to-noise ratio" matrix are derived under both sample rich and sample poor scenarios. We show that when the number of data samples is smaller than the dimension of the data channel, the eigenvalues of the signal-to-noise ratio matrix become defective and the empirical J-divergence does not measure the actual separation between the two hypotheses. This is more critical when using a kernel version of the detector owing to the high dimensionality of the mapped feature space. A numerical example is also presented which illuminates these properties of the Gauss-Gauss detector under sample rich and sample poor conditions.

James Derek Tucker  
Department of Electrical and Computer Engineering  
Colorado State University  
Fort Collins, CO 80523  
Spring 2009

# ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Mahmood Azimi, for his entire support and guidance throughout the process of my Masters. He has taught me the value of well-developed research and the ability to explain my thoughts. His assistance and time is greatly appreciated throughout my education.

I would also like to thank my committee members, Dr. Rockey J. Luo and Dr. F. Jay Breidt, for their time and assistance.

I'd also like to thank the Naval Surface Warfare Center Panama City for providing the funding and the data used in this project. This project was funded by the Office of Naval Research (ONR-321OE) under contract numbers N61331-06-C0027 and N00014-08-1-0142. Without this funding from the ONR, this work would never have been completed. I would also like to thank Dr. Gerry Dobeck, Dr. Jason Stack, and Tory Cobb at NSWC-PC for their continuous support, suggestions, and help throughout this project. They were willing to help anytime and have provided me with valuable insight to the problem of detection of targets from sonar imagery

Special thanks goes to my colleagues in the Signal and Image Processing Lab. They have provided a great environment for discussing work, help, and even times providing that needed break. Thanks to Nick R., Neil, Mike M., Nick K., Mike K., and Jon.

Finally, I would like to thank my wife and my family. They have always believed in me and provided the support and guidance to pursue my dreams. Without them I would have never gotten this far.

*To my parents and my sister, thank you for love and support throughout my education.  
And to my wife Nichole, you have been my support and my love throughout this  
process.*

# TABLE OF CONTENTS

<b>SIGNATURE</b> . . . . .	<b>ii</b>
<b>ABSTRACT OF THESIS</b> . . . . .	<b>iii</b>
<b>ACKNOWLEDGEMENTS</b> . . . . .	<b>vi</b>
<b>DEDICATION</b> . . . . .	<b>vii</b>
<b>1 INTRODUCTION</b> . . . . .	<b>1</b>
1.1 Problem Statement and Motivations . . . . .	1
1.2 Literature Review . . . . .	5
1.3 Research Objectives . . . . .	10
1.4 Organization of the Thesis . . . . .	13
<b>2 DATA DESCRIPTION AND PRE-PROCESSING METHODS</b> .	<b>15</b>
2.1 Introduction . . . . .	15
2.2 Experimental Setup . . . . .	16
2.2.1 Sonar8 Data Set . . . . .	17
2.2.2 Multi-Platform Sonar Data Set . . . . .	19
2.3 Preprocessing Methods . . . . .	20
2.3.1 Image Normalization [1] . . . . .	21
2.3.2 Image Partitioning and CCA-Channelization . . . . .	30

2.4	Conclusion . . . . .	34
<b>3</b>	<b>BINARY HYPOTHESIS DETECTION METHODS . . . . .</b>	<b>36</b>
3.1	Introduction . . . . .	36
3.2	Bayesian Detection Framework . . . . .	38
3.2.1	Neyman-Pearson Detection . . . . .	40
3.2.2	Gauss-Gauss Signal Detection . . . . .	41
3.3	Gauss-Gauss Signal Detection in CCA Framework . . . . .	44
3.4	Conclusion . . . . .	47
<b>4</b>	<b>DETECTION OF TARGETS IN SONAR IMAGERY . . . . .</b>	<b>49</b>
4.1	Introduction . . . . .	49
4.2	Composite Channel Gauss-Gauss Detection and Formulation [2] . . . . .	51
4.3	Single-Sonar Implementation and Results . . . . .	56
4.3.1	System Overview . . . . .	57
4.3.2	Results and Observations . . . . .	58
4.4	Dual Disparate Sensor Implementation and Results . . . . .	61
4.4.1	System Overview . . . . .	63
4.4.2	Results and Observations . . . . .	64
4.5	Conclusion . . . . .	66
<b>5</b>	<b>MULTI-PLATFORM DISTRIBUTED DETECTION IN SONAR</b>	

<b>IMAGERY</b> . . . . .	<b>69</b>
5.1 Introduction . . . . .	69
5.2 Distributed Detection . . . . .	70
5.3 Multi-Platform Distributed Implementation and Results . . . . .	74
5.3.1 System Overview . . . . .	75
5.3.2 Results and Observations . . . . .	77
5.4 Conclusion . . . . .	82
<b>6 EFFECT OF SAMPLE SUPPORT ON LINEAR AND KERNEL GAUSS-GAUSS DETECTION</b> . . . . .	<b>84</b>
6.1 Introduction . . . . .	84
6.2 Gauss-Gauss Detection and Data Sample Support . . . . .	86
6.2.1 Sample Rich Case ( $m < M$ ) . . . . .	88
6.2.2 Sample Poor Case ( $m > M$ ) . . . . .	89
6.3 Kernel Gauss-Gauss Detection . . . . .	91
6.4 Results and Observations . . . . .	94
6.5 Conclusion . . . . .	98
<b>7 CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK</b>	<b>100</b>
7.1 Conclusions and Discussions . . . . .	100
7.2 Future Work . . . . .	105

APPENDIX A — CANONICAL CORRELATION ANALYSIS RE-  
VIEW . . . . . 117

# LIST OF FIGURES

1.1	Single Sensor Approach. . . . .	2
1.2	Multiple Disparate Sensor Application. . . . .	4
2.1	Formation of a Target in a Sonar Image. . . . .	18
2.2	Illustration of SFBF Path Selection. . . . .	23
2.3	Illustration of the Normalization Method. . . . .	24
2.4	Flow Chart of SFBF Normalization Method. . . . .	24
2.5	Original Side-Scan Sonar Image. . . . .	28
2.6	SFBF-Normalized Side-Scan Sonar Image. . . . .	29
2.7	Channelization for CCA for the Single Sensor Case. . . . .	32
2.8	Channelization for CCA for the Dual Disparate Sensor Case. . . . .	33
3.1	CCD used as an optimal detector. . . . .	45
4.1	Coherence-based Detector for Composite Channel Hypothesis Case. . . . .	52
4.2	Block Diagram of the Single-Sensor CCA-based Detection Method. . . . .	56
4.3	Plot of Canonical Correlations for Target and Background for Sonar8. . . . .	59
4.4	Histogram of Example Log-Likelihood Values for Target and Background for Sonar8. . . . .	60
4.5	ROC Curves for the Sonar8 Data Set. . . . .	61

4.6	Coherent-Band detection and feature extraction from Dual Disparate Sonar Platform. . . . .	62
4.7	Plot of Canonical Correlations for Target and Background for HF-BB.	65
4.8	Histogram of Example Log-Likelihood Values for Target and Background for HF-BB. . . . .	66
4.9	ROC Curve for the HF-BB Detector. . . . .	67
5.1	Block Diagram of the Distributed Detection Problem. . . . .	71
5.2	Block Diagram of the Multi-platform Detection System. . . . .	75
5.3	Plot of Canonical Correlations for Target and Background for both Detectors. . . . .	78
5.4	Histogram of Example Log-Likelihood Values for Target and Background for both Detectors. . . . .	80
5.5	ROC Curves for the Individual Detectors and the Fusion . . . . .	81
6.1	Eigenvalue Gap versus Number of Samples. . . . .	95
6.2	J-divergence versus Number of Samples. . . . .	95
6.3	Probability of Detection versus Number of Samples. . . . .	96
6.4	Probability of False Alarm versus Number of Samples. . . . .	96
A-1	Transformation from standard coordinates $\mathbf{x}$ and $\mathbf{y}$ to canonical coordinates $\mathbf{u}$ and $\mathbf{v}$ . . . . .	120

# LIST OF TABLES

2.1	Forward Left to Right Filter Equations in Order of Execution. . . . .	25
2.2	Forward Right to Left Filter Equations in Order of Execution. . . . .	26

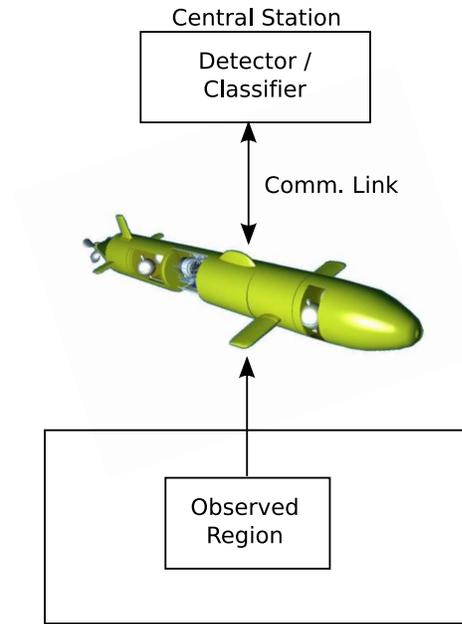
# CHAPTER 1

## INTRODUCTION

### 1.1 Problem Statement and Motivations

The problem of underwater object detection and classification in sonar imagery has attracted a substantial amount of attention [1] - [16]. This problem is complicated due to various factors such as variations in operating and environmental conditions, presence of spatially varying clutter, variations in target shapes, compositions and orientation. Moreover, bottom features such as coral reefs, sand formations, and vegetation may totally obscure a target or confuse the detection process. Consequently, a robust detection system should be able to quantify changes between the returns from the bottom and any target activity in sonar images, while at the same time extract useful features for subsequent classification. Thus, a system designed without the need to perform separate detection and feature extraction is highly desirable.

Normally, a single sensor (sonar, lidar, etc.) is used to detect and classify the objects based upon observations taken from the environment. From these observations, the sensor will either make a local decision and transmit it to a central station or record the entire sonar image for post mission analysis (PMA) at the central station (see Figure 1.1). The issues faced with detection based upon one sensor is that the detection process is limited to only one field of view. This makes the detection of weak targets particularly challenging. Moreover, the structure of targets within an image vary as a function of aspect, grazing angle, and range from the sonar which makes detection difficult, especially if the target is in a disadvantaged position in relationship to the sensor, e.g. partially obscured targets. Therefore, any improvement of the detection results is limited due the limited amount of data and observations



**Figure 1.1:** Single Sensor Approach.

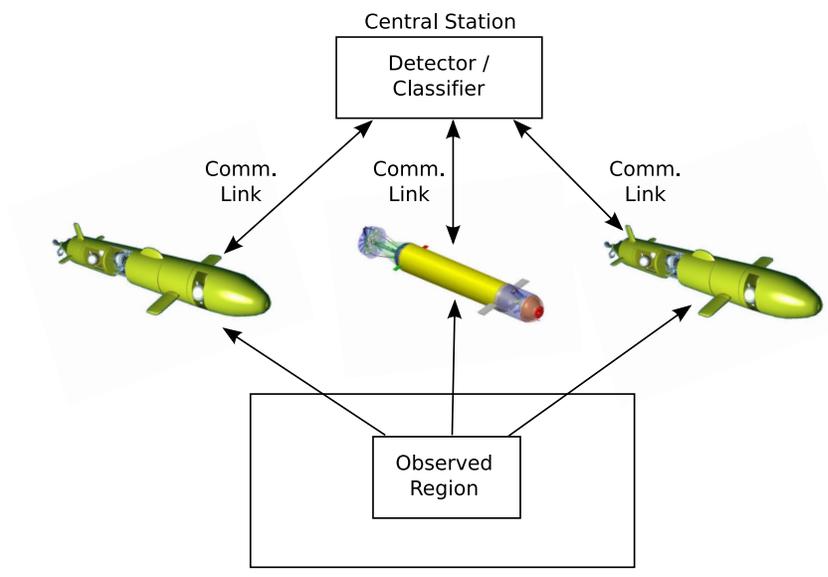
from the environment.

Distributed sensor networks offer and new a promising solution to overcome the shortcomings of the single sensor situations. The use of multiple sensors allows for significantly better capture of the target characteristics due to the fact that the targets are viewed from different aspects, grazing angles, ranges, frequencies, and sensor modalities. In a surveillance area there could be multiple autonomous underwater vehicles (AUV's) each equipped with a wide variety of sensors including different types of sonar, magnetics, or electro-optical systems or a single AUV equipped with multiple sensors. Preliminary detection, feature extraction, and object classification can be performed based upon the data collected using every sensory system on one or multiple vehicles. A final decision-making usually takes places at the central station, either in the standard PMA method or real-time network-centric sensor analysis (NSA) using some type of decision-level or feature-level fusion. However, due to the increase in the number of sensor the amount of bandwidth is increased by the number of sensors requiring careful design of the detection method.

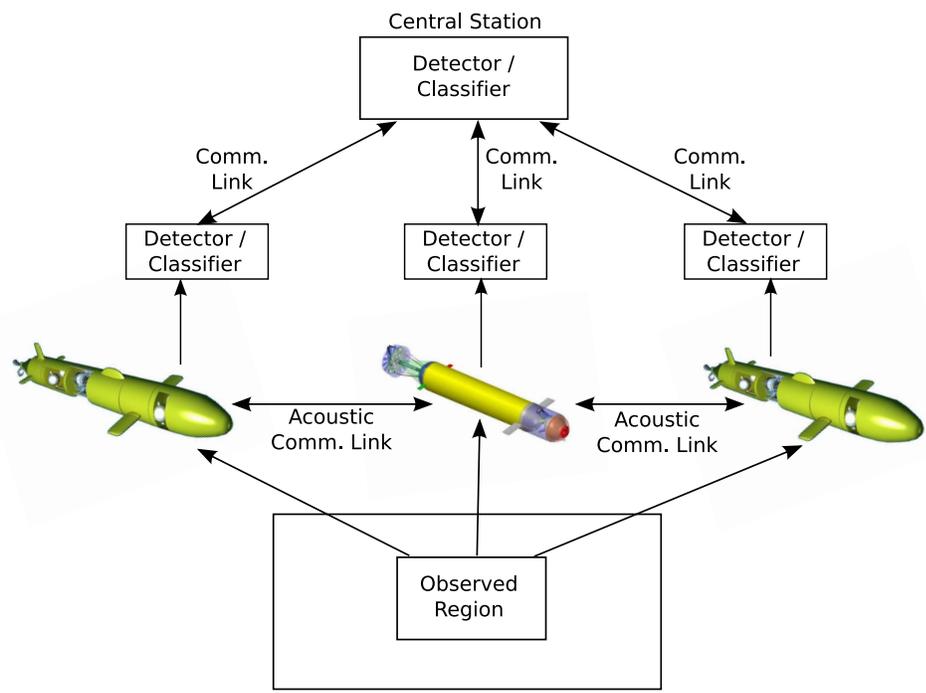
In the PMA situation each sensor transmits the entire sonar image to the central station, where all the low-level and high-level processing and fusion based upon the transmitted images (see Figure 1.2(a)) take place. Since preliminary decisions and feature extraction are typically carried out based upon independent sensor data, the fusion may only have access to partial or incomplete information resulting in a loss of detection performance.

In real-time NSA situations, multi-sensor collaboration among distributed sensor reduces the uncertainty in decision making. Each sensor makes a local decision based upon its observation and shares this observation and/or preliminary decision to other sensors over a communication link. Each sensor then makes a final decision based upon several observations and/or preliminary decisions received and then sends its decision to the central station (see Figure 1.2(b)). However, this situation requires careful consideration of the local (sensor-level) computational requirements and limited network communication between the sensor platforms and central station. Therefore, a careful selection of what attributes of the observation should be sent due to the very low bandwidth of the acoustic communication link is needed.

To allow collaborative decision-making among multiple sonar platforms, it is essential to detect and further scrutinize the information bearing parts of the data collected by the various sensory systems. This involves detecting, isolating, and representing; in terms of some pertinent attributes, the *coherent*, or mutual information among one or multiple data sets. This is an extremely challenging problem due the disparate nature of the problem and therefore to develop such a solution, new methodologies are needed to: (a) collaboratively detect and agree on threats occurring within the field of view of the sensors, (b) perform feature extraction to capture common target attributes from multiple sensor platforms, (c) perform object classification and identification, (d) and finally develop a single integrated target assessment picture based upon the detected, localized and classified targets from one or multiple disparate



(a) PMA



(b) NSA

**Figure 1.2:** Multiple Disparate Sensor Application.

sensors.

In this work we attempt to address these problems by designing a coherence-based detector where the coherent information is found between multiple channels and is used to detect objects from sonar imagery captured from one or multiple sonar platforms with a high degree of confidence.

## 1.2 Literature Review

Considerable research has been devoted to the development of different detector and classification methodologies to detect and classify underwater objects from sonar imagery. Dobeck [4], [5] utilized a nonlinear matched filter to identify mine-size regions that match the target template in a side-scan sonar image. For each detected region, several features were extracted based on the size, shape, and strength of the target template. A stepwise feature selection process was then used to determine the subset of features that maximizes the probability of detection and classification. A k-nearest neighbor and an optimal discrimination filter classifier were used to classify each feature vector and the decisions of the two classifiers were fused to generate the final decision.

In [6], a method was proposed that first median filters the sonar image to reduce the speckle noise present in the image. The image was then split into overlapping range segments where the pixels in each segment were adaptively thresholded. The threshold was determined from cumulative distribution function (CDF) formed from a training set. The purpose of the thresholding is to identify the target structure in the processed segments. Geometric features were then extracted from contiguous target structure regions within the segment. The regions within the segment correspond to the highlight and shadow structures of the target. Classification of each region as target or non-target was done through a multi-level weighted scoring-based classification system. The algorithm was extended in [17] to include a normalization

algorithm to reduce the image speckle and reduce image artifacts. A quantization was also applied to reduce the dynamic range of the pixel intensities [6]. The addition of the normalization algorithm had an overall reduction in the total number of false alarm for the same probability of detection. The reason for reduction in the number of false alarms is that the normalization step reduces the variability of the background even for cluttered environments. Moreover, the normalization algorithm provides better target definition which improves the shape features that are applied to the classifier.

An adaptive clutter filter detector was presented in [3] which exploits the difference in correlation characteristics between clutter and targets. After detection by the adaptive clutter filter, features were extracted from the detection regions and then orthogonalized. The orthogonalization is done by forming the overall scatter matrix which is the sum of the individual scatter matrices for each of the individual classes. A transformation matrix is then found such that when applied to the overall data scatter matrix, the eigenvalues of the scatter matrix are unity. Classification is then performed on the orthogonalized feature set using an optimal Bayesian classifier [18].

In [11], [19], the algorithm described in [3] was applied to three different sonar images varying in frequency and bandwidth. The classification on each image was done using a multistage classification approach, which entails a repeated application of a classifier. During the training stage, it is determined how many times to apply the classifier and the optimal subset of the feature vector that is used at each application. Each stage of the classifier results in a reduction in the number false alarms. The final classification decision is made by a fusion of the three classification results from the three different sonar images. This work can be regarded as the first attempt to use disparate sonars where the disparateness is in the operating frequency of the sonar. However, in this application the detector and classifier process each image individually and do not use the information contained in the three images simultaneously

to make the detection and classification call. This type of decision-making, which is based upon individual sensory data, typically leads to incomplete, degraded or biased decisions. However, when the information from the individual sensors is used collaboratively and simultaneously in the decision-making a more complete decision about the observation can be made.

Chandran in [9] presented the use of a matched filter designed to capture the target structure. Higher order spectra were extracted from the phase of the Fourier transform as the feature set to classify objects. A k-nearest neighbor classifier [18], a minimum distance classifier [18], and a threshold-based classifier where the threshold is determined from the minimum and maximum values of a feature obtained over all classes were used. The outputs of the three classifiers were fused to yield a final decision. However, the use of a matched filter is not ideal, as targets can vary greatly in shape, composition, and orientation in the sonar images and hence leading to misdetections.

Atallah uses a scale-saliency [20] feature approach for the detection of objects in side-scan sonar images. The scale-salient features were found by first taking each pixel and calculating the Shannon Entropy [21] over a set of circular windows around the current pixel where the radius (scale) is changed. Next, the scales are selected such that the entropy is maximized and to select the most *salient* ones. The statistics of the local window are used to weigh the entropy for each of the peaks. This is known as the weighted Shannon Entropy [22] and from which the scale and final saliency feature is chosen for the current pixel. After each pixel's scale and saliency were determined, a clustering algorithm was used to group together pixels which possess similar salient features and scales. The clustering algorithm groups together similar neighboring pixels to form individual salient regions. Each region was assigned a measure which consisted of the mean of the scales within that region. Using a threshold on this measure, objects were detected for regions where the mean of the scales fell above

the predetermined threshold. Although the features are formed from specific pixel attributes, this method when used in a large region of clutter or texture can produce a high amount false alarms due to the grouping of regions with high scales and entropy.

A Markov random field (MRF) model approach was studied in [23] where the required model parameters were estimated from the original image. By using the *a priori* spatial information on the physical size and geometric structure of possible targets in the field, the MRF model [24] directly segments the image into regions of highlight, shadow, and background. After detection, features were extracted using a cooperating statistical snake model [25] which extracts the boundary of the highlight and shadow from the detected regions. The snake again utilizes available *a priori* information about the spatial relationship between the highlight and shadow structures. After the features are extracted they are classified using a Bayesian classifier [18].

The one pitfall in using a model-based approach in object detection is knowing the *a priori* information necessary to construct the model. This can be extremely difficult due to changing seafloor conditions and target compositions and the large number of non-mine-like objects, either man-made or natural, that may be encountered. Thus, if any of the prior information is incomplete or incorrect the model generated from the data will suffer and consequently the detection rate will fall and the false alarm rate will increase. Dura in [26] developed an active learning algorithm based on kernel type classifiers to overcome this pitfall of knowing the *a priori* information. A set of the best basis functions was determined from the observed data and the number of required basis functions is determined adaptively by an algorithm using the Fisher Information matrix [18] associated with the observed data and the set of basis functions. Basis functions are added to the basis set until the information gain is no longer deemed significant. Once the basis set is defined the associated model weights are determined for a kernel-based classifier. The optimal training data was determined adaptively in-situ via information-theoretic metrics by iteratively adding

cases until the information gain is below a prescribed threshold.

Until recently, one detection and feature extraction method that was not explored for sonar imagery was CCA [27] - [29]. This method does not require any *a priori* information about the targets or the environment that contains them. This method recently has shown great promise in underwater target classification problems using sonar backscatter data [30]. The canonical coordinate decomposition method determines linear dependence [27] or coherence between two data channels. This method not only determines the amount of dependence (or independence) between two data channels (e.g. two sonar pings with certain ping separation) but also extracts, via the canonical coordinates, a subset of the most coherent features for classification purposes. CCA allows one to quantify the changes between the returns from the bottom and when target activities are present and at the same time extract useful features for target classification without the need to perform separate detection and anomaly feature extraction. The CCA method has shown great promise in underwater target classification problems using sonar backscatter using data collected by the buried object scanning sonar (BOSS) system [30] - [33]. The work in these references presented a multi-ping classification system that extracts coherence-based features from blocks of range cells of a time series of two sonar returns with single ping separation. These coherence patterns were shown to be different for pairs of pings that contain mine-like objects than those that contain non-mine-like objects. The canonical correlations that capture the coherence patterns [27] were shown to have high discriminatory power for both detection and classification. Another study, [34] extends this coherence analysis to the frequency domain by measuring coherence between the same frequency band in two sonar pings. Comparing to the time domain coherence-based features, these features provided substantially better results on the BOSS data sets as well as the ability to offer a more rigorous way of generating acoustic-color for possible target identification from multiple sonar pings.

Recently in [14] - [15] we have demonstrated the use of CCA in the detection and classification of objects from sonar imagery. These results showed that by using the canonical correlations extracted from consecutive columns of an ROI within a sonar image we can perform target detection. From the detected ROI's the corresponding canonical correlations can also be used for classification using a single classifier. The initial results motivated us to develop the methods presented in this thesis. The specific research objectives are described next.

### 1.3 Research Objectives

The goal of this work is to develop and test an efficient and robust coherence-based detection system for sonar imagery that maintains good underwater target detection performance in varying operating and environmental conditions. Specifically, we would like to develop and test a detection system that can either be applied to single sonar platform or to dual (or multiple) disparate sonar platform cases using the data provided by the Naval Surface Warfare Center, Panama City (NSWC-PC). For the single sensor application, our detection hypothesis is that for a ROI within a sonar image, a target will have a higher level of coherence between columns in the ROI than that if background only was present. For the dual disparate sensor applications, our detection hypothesis is that the presence of a target in a co-registered pair of ROI's will have a higher level of coherence than if the pair contained only background. As stated before, the CCA method [27] - [29] provides an excellent framework for quantifying changes between the returns from the bottom and the returns from a target by determining the linear dependence (or coherence) between two data channels (e.g. two columns in a ROI). This method not only determines the amount of dependence between two data channels, but also extracts, via the canonical correlations [27] - [29], a subset of the most coherent features for classification purposes. This detection and feature extraction capability can be done without the need to perform separate

detection and anomaly feature extraction.

Unfortunately, the standard Gauss-Gauss detector that was cast in the CCA framework [29] cannot directly be applied to this problem, owing to the fact that in both applications (single platform or multiple-platform) both hypotheses involve two channels. That is, for the null hypothesis,  $H_0$ , the ROI(s) cover background only while  $H_1$  corresponds to the case where the ROI(s) contain target (signal) and background. Specifically, for the single sensor application the two columns in an ROI, which form the two channels for the CCA framework, will both contain background under  $H_0$  or target plus background under  $H_1$ . For the dual disparate sensor case, a pair of blocks from a pair of co-registered ROI's, which again are the two channels for the CCA framework, will both contain background under  $H_0$  or target plus background under  $H_1$ . With the change in the two hypotheses the expressions for the log-likelihood and J-divergence in [29] no longer hold and will have to be reformulated.

In this thesis, a development of the composite two-channel Gauss-Gauss detector which overcomes the above-mentioned shortcoming will be presented. Moreover, the development in [29] uses the assumption that one of the CCA channels is always signal only and the other channel is the observation. This does not lend itself well to our detection problem or the development of the composite two-channel detector. The work presented here will develop the log-likelihood and J-divergence for the composite two-channel hypothesis testing and relate the formulations to the standard Gauss-Gauss detector presented in [35]. From this development we then cast the composite two-channel Gauss-Gauss detector it in the CCA framework where the assumption is not made that one of the CCA channels is signal only.

After the development of the composite two-channel coherence-based Gauss-Gauss detector, two different applications of the detector will be studied, one for the single sensor platform case and one for the dual disparate sensor platform case. A comprehensive analysis of the proposed applications will be carried out on two sonar

data sets. These data sets will correspond to different environmental and operating conditions, mine-like and non-mine-like objects and different clutter density. More specifically, the two data sets will correspond to the two applications studied in this thesis. One data set contains single sensor sonar imagery captured from a single side-scan sonar and will be applied to the first application. For the dual disparate sensor application, there is a multi-platform sonar data set that contains multiple co-registered imagery captured using sonars with different resolution and frequency characteristics.

Due to the nature of the problem there are few necessary pre-processing steps that will have to be evaluated. First, due to the fact that most sonar images contain noise, speckle, and artifacts from the image formation process, a normalization process for the high frequency sonar images will be used. Second, due to the application differences of the single sensor case and the dual disparate sensor case, the methods for channelization, vectorization, and averaging for the CCA process is explored. This is especially true for the differences in the channelization techniques for the two cases, where the channels need to be chosen to optimally extract the coherent information. Finally, for the dual disparate sensor case the system will have to take into account differences in image and target sizes due to differences in the sensors operating frequency, range, altitude, etc.

The dual disparate application will then be extended to the distributed detection scenarios where the coherence-based detector will be used as multiple local decision makers. Each of the decision makers will then transmit their decisions to a fusion center where a final decision will be made. The motivation behind this application is that it accounts for (a) limited communication bandwidth for communicating essential information among multiple disparate AUV's and the mother ship; (b) computational limitations of the DSP boards on each AUV platform; and (c) near real-time distributed detection. Moreover, by using multiple detectors one can greatly reduce the

false alarm rate and increase the detection rate by using more observations from the environment. A new fusion rule for the fusion center will be developed in this thesis and the results on the multi-platform sonar data set will be benchmarked against the single dual disparate sensor case.

The effectiveness of the developed algorithms will be analyzed for various background difficulty and target compositions. The results of applying these systems to two different sonar imagery data sets are evaluated and thoroughly discussed in order to determine the effectiveness of the proposed coherence-based detector for the detection of underwater objects from side-scan sonar imagery.

Finally, a study will be conducted on the sample-support of the Gauss-Gauss detector. When implementing the detector the covariance matrices under the two-hypotheses have to be estimated from a limited number of samples drawn from the corresponding observations. A study of the effectiveness of the detector under two case will be explored, namely the sample rich and sample poor cases. In the sample rich case, the number of samples is greater than the data vector dimension while for the sample poor case the opposite is true. Although the sample poor scenarios may not occur frequently for the linear detector, they do indeed happen in the kernel nonlinear version of this detector. This is due to the fact that the dimension of the mapped data in the high dimensional feature space is typically much higher than the sample support [36] - [38]. From the study of the sample poor scenarios the kernel version of the Gauss-Gauss detector will then be developed.

## **1.4 Organization of the Thesis**

This thesis is organized as follows: Chapter 2 introduces the two sonar imagery data sets used in this study together with a description of the different properties of the data sets. An outline of the preprocessing methods and image normalization

procedure is also presented. In Chapter 3, a detailed review of Bayesian and Gauss-Gauss detection is provided. The development of the Gauss-Gauss detector cast in the CCA framework is also reviewed. The development of the two-channel coherence-based detector for sonar imagery is then given in Chapter 4. In this chapter, the two applications of the detector are examined. The first implements the detector on the single-sensor case and examines its attributes and properties with results provided using the single-sonar data set. The second implementation is on the dual disparate sonar case and again examines its attributes and properties with results provided using a multi-platform sonar data set. In Chapter 5, the development of the distributed detection system and the corresponding decision fusion rule is provided along with results using the dual disparate detector as the local decision maker and on the multi-platform sonar data set. Chapter 6 studies the sample support of the empirical Gauss-Gauss detector and introduces the development of the kernel Gauss-Gauss detector. In order to develop the kernel version of the detector an analysis of the sample support of the linear Gauss-Gauss detector is examined. Finally, Chapter 7 concludes the studies carried out in this research and discusses the goals for future work.

# CHAPTER 2

## DATA DESCRIPTION AND PRE-PROCESSING METHODS

### 2.1 Introduction

This chapter will discuss the sonar imagery data sets that are used for the experiments and studies presented in this thesis. Two types of data sets are used in this study, with both presenting different properties and challenges for target detection. Analyzing the different properties of each data set is an important issue, as understanding the type of sonar imagery data and how images are formed can provide valuable information on how to effectively implement a detection method. The first data set consists of envelope data [39], the magnitude of the complex data, which is the output of a beamforming process once all the sonar returns have been processed. This data set was collected using only one sonar system which was a high-resolution high frequency side scan sonar. The second data set corresponds to a disparate multi-sonar data set consisting of one high frequency sonar and three broadband sonars. The images contain complex data, which is the direct output of the beamformer after all the sonar returns have been processed and co-registered over the same region of the target field. For this data set disparateness of the sonar systems is with respect to operating frequencies and beamwidth rather than platform locations.

In order to prepare the sonar imagery for detection, a few preprocessing procedures must be applied to the images. First, a normalization method is applied to remove sonar artifacts, reduce clutter, and enhance the target signature. The purpose of the normalization is to reduce the variability of the local mean throughout the image in

order to use it as a reference level so that the highlight and shadow of the target can be more easily identified. The serpentine forward-backward filter (SFBF) normalizer [1] uses a second order digital filter and attempts to select a path in which the filter output best follows the original image, i.e. the filter path follows the direction of least change in the image. The image is filtered in the forward and backward directions to estimate the local mean on both sides of a pixel. The local mean estimate that is nearest to the original image value is selected to normalize that pixel. After the image is normalized it is then partitioned into overlapping ROI's to prepare the image for the CCA-based detector. Each ROI is further partitioned into either columns for the single sensor approach or pairs of blocks for the dual disparate sensor approach that form the realizations for the two-channels in CCA. The covariance matrices in the CCA are formed by averaging over all columns (or blocks) within the ROI's .

The outline of this chapter is as follows. Section 2.2 gives an overview of the experimental setup by reviewing the two data sets used in this study together with their properties. Section 2.3 discusses the preprocessing, namely the image normalization process that is used to reduce background clutter and enhance the target signature and the methods to channel, vector, and average the data for CCA for the different applications. Finally, conclusions are made in Section 2.4.

## 2.2 Experimental Setup

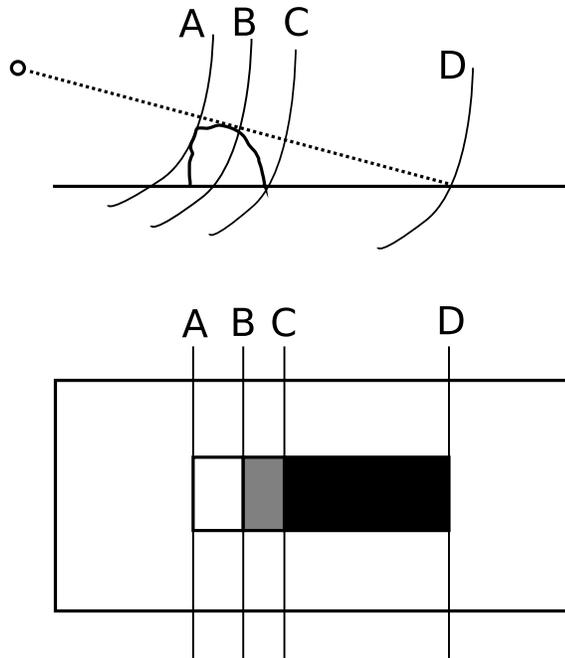
In this thesis, two separate data sets are used to test the effectiveness of the developed detector and analyze the results. Although the detector developed in this thesis is used on both data sets, the manner in which it is implemented differs for each data set due to the variations in sonar set up (e.g. single sensor or disparate sensor cases), frequency, bandwidth, image size, and resolution of the sonar used in the respective data collection. This will be discussed in more detail in the next section for the two different data sets that are referred to as *Sonar8* and *multi-platform sonar* data sets

throughout the rest of this thesis.

### 2.2.1 Sonar8 Data Set

The first data set, referred to as *Sonar8*, was provided by the NSWC in Panama City, FL. This data set consists of a high resolution side-scan sonar that carries only one type of sonar. More information on high-resolution side-scan sonar can be found in [39], [40]. The data set contains only the envelope data, i.e. the values of the pixels are the magnitude of the complex image. The complex data (magnitude and phase) are generated at the output of a coherent processor, in this case the k-space or wavenumber beamformer [41], [42]. Each impinging sound wave on the receiver array elements of the sonar is converted to magnitude and phase. The delay and sum beamforming algorithm [43] attempts to coherently combine the sound waves in a way that resolves the echo returns into a complex-valued pixel. More specifically, the k-space or wavenumber algorithm [41], [42] computes the 2-D Fourier transform of the raw or range-compressed sonar data in the delay-time/aperture domain. This converts the data into the spatial frequency/wavenumber  $(\omega, k)$ -domain where it is multiplied by the power spectrum of the transmitted wavefront. A change of variables is done by Stolt interpolation [44]. This change of variables maps the frequency/wavenumber  $(\omega, k)$ -domain into the wavenumber domain  $(k_x, k_y)$ . The inverse 2-D Fourier transform is then taken of the mapped data to form the complex image. For more information on the k-space/wavenumber beamformer the reader is referred to [42].

In order to better understand the difficulty of underwater object detection from sonar imagery it is important to understand the formation of a target signature in a sonar image. Figure 2.1 shows how the signature of a target is formed in a sonar image. In this figure, region A-B corresponds to the *highlight* or a strong sonar return off of the object, region B-C is known as the *dead zone* where no sonar return can



**Figure 2.1:** Formation of a Target in a Sonar Image.

occur due to the return being blocked by the object, and region C-D is known as the *shadow*. The size of each of these regions greatly depends on the type of target, range, height, aspect, and grazing angle of the sonar with respect to the object. More specifically, the signature of a particular target can vary greatly as a function of range from the sonar. As the target is further out in range from the sonar, the shadow region (C-D) will become longer. Additionally, the overall signature becomes less defined with less definition between the highlight, dead zone, and shadow regions. This creates a particular problem for most detection methods which rely on detecting the specific structure of the target. The coherence-based detector developed in this thesis (see Chapter 4) provides a method to detect targets even at far range with high probability as it looks for coherence and not just the specific target structure.

The Sonar8 data set contains 137 images containing 286 targets with most of the images containing multiple targets. The data set was broken up into easy, medium, and hard cases depending on the difficulty of the background clutter and bottom types. Easy cases are considered to have low background variation and an overall

smooth bottom with targets that are easily identifiable by a skilled operator. The medium cases contain background clutter and more difficult bottom conditions. However, the targets are still somewhat discernible to a skilled operator with some effort. Finally, the hard cases are those where it is difficult to detect and distinguish the targets from a visual inspection due to a high variability of background clutter and very difficult bottom conditions. The hard cases also contain the cases where the targets are obscured by dense clutter with none of the targets easily detectable by the naked eye. The data set was separated into these three classes based on a visual inspection of where the targets are located and whether or not variation in the background and high density of clutter are present.

To help enhance the ability to detect objects in this data set a normalization preprocessing is applied to each image to help distinguish the target’s highlight and shadow signature from the bottom and artifacts present in the image. A review of this normalization process is given in Section 2.3.1.

### **2.2.2 Multi-Platform Sonar Data Set**

This multi-platform sonar data set was also provided by the NSW-C-PC. The data set contains a high frequency (high resolution) side-looking sonar image over the target field and three broadband sonar images co-registered over the same region. The disparateness of this data set comes not from the location of the sonar platforms, but from the difference in operating frequencies and beamwidth of the sensors. More information on high resolution side looking and broadband sonar can be found in [39], [40] and [45], [46], respectively. The images in this data set are also complex corresponding to the output of the (k-space / wavenumber) beamformer [41], [42] which was described earlier. The benefit in using the complex data rather than just the envelope data as in the Sonar8 data set is that it provides more information by providing both the magnitude and the phase information. As mentioned before, the

sonar images captured by different sensor systems are co-registered so that the pixels correspond to the same section on the seafloor. This is important due to the fact that each sensor is operating at a different frequency and beamwidth and has different spatial resolution hence providing a different size sonar image.

This data set contains 59 images from all 4 sonar systems and contains 53 targets with some of the images containing more than one target. The benefit in using a data set containing four disparate sensors each looking at the same piece of the ocean floor is that it helps reduce false alarm in the detection process due to its multi-look nature. This is mainly due to the fact that the high frequency sonar provides higher spatial resolution and better ability to capture target details and characteristics. In the high frequency sonar, the highlight, dead zone, and shadow structures will have good definitions with each having a definite shape and adequate number of pixels. Nonetheless, these structures will also be prominent for structured clutter as well (e.g. coral reefs and sand ripple). In a broadband sonar, however, the target structure is not as well-defined. In other words, a target will primarily show up as a highlight with no prominent dead zone or shadow structure. Nonetheless, the signature of the clutter is suppressed hence leading to substantially lower number of detected contacts. Therefore, the use of multiple disparate sonar systems allows one to exploit a high resolution sonar with good target definition while taking advantage of the clutter suppression ability of a low resolution broadband sonar co-registered over the same region to provide potentially much better detection performance comparing to those of the single sonar cases.

## 2.3 Preprocessing Methods

In order to apply the detection scheme to sonar imagery, three preprocessing methods are used to prepare the data. First, the sonar image is normalized using a serpentine forward-backward filter (SFBF) [1] to help reduce clutter and enhance the signature

of the target. The normalized image is then partitioned into overlapping ROI's for processing. Each ROI is then subsequently partitioned into blocks that form the realization for the two-channels in the CCA method for extracting the canonical coordinates and correlations. The partitioning is done such that each block is reshaped into a column vector, where each vector makes up a column in the data matrix for the corresponding ROI. A detailed description of these methods is provided in the following sections.

### **2.3.1 Image Normalization [1]**

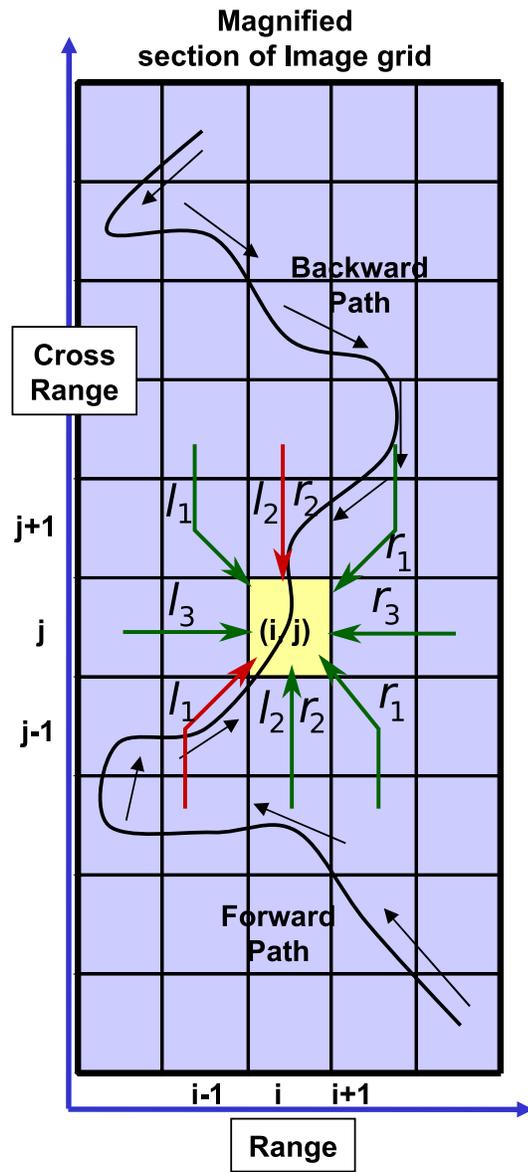
In sonar imagery the target signature consists of bright and dark intensity pixels that make up three distinct regions described earlier (see Figure 2.1). The challenge in the target detection is to reliably distinguish the target's highlight and shadow signatures in a highly cluttered background region within the image. The purpose of image normalization is to reduce the variability of the local mean throughout the image to a relatively constant level so that each highlight and shadow pixel can be more easily discerned. As mentioned before, the normalization method used on the data sets described previously employs a SFBF [1] method that addresses several concerns in image normalization process. First, the normalizer should function properly throughout the image as the presence and signature characteristics of the targets are not known. Second, the normalization should not lead to any false alarms around the edge or between regions with very different pixel intensity values. Third, the process should have a low computational needs.

The SFBF [1] uses second-order, recursive 2-D digital filters to generate two estimates of the local background at each processed pixel. The first estimate approximates the local mean on one side of the pixel, and the second estimate approximates the local mean on the other side of the pixel. The first estimate is generated as a delayed output of the filter, in terms of location in the image, running along the

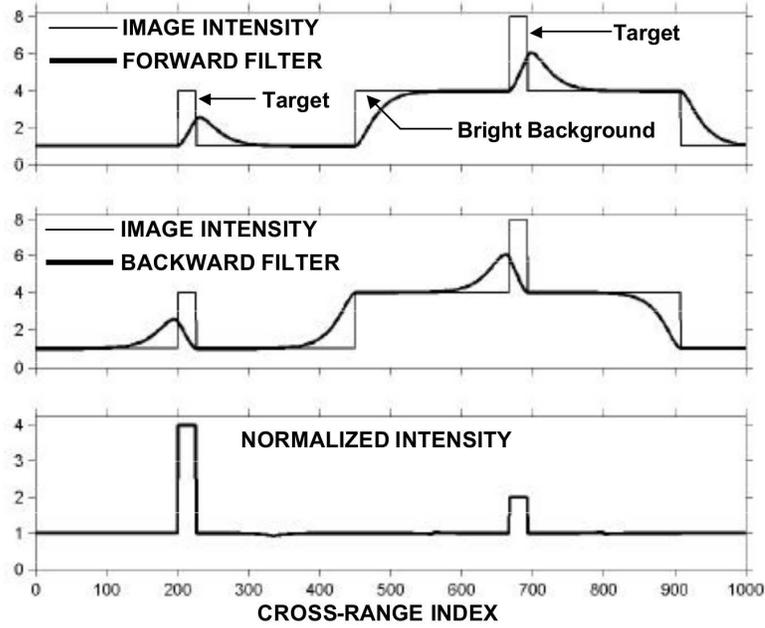
cross-range dimension in the image proceeding in the forward direction. The delay is selected to be the size of the largest expected target in the cross-range direction. This value is assumed to be known based upon average target size in the field and the resolution of the sonar. This delay is introduced so that the estimate of the local mean at a given pixel location in the image is not compromised by a potential target that may be at the same location. Similarly, the second estimate is generated as a delayed output of the filter running along the cross-range dimension in the backward direction. Of the two local estimates, from the forward and backward filters, the one that is nearest the original image value is selected to normalize that pixel. Operating in the forward and backward directions allows the target edges to be preserved while eliminating pixel variations in the background.

The SFBF filter attempts to select a path along which the filter output best follows the original image. This idea is presented in Figure 2.2. Specifically, the path is generated recursively by extending the path's latest end point to one that belongs to a subset of its neighbors whose intensity is most near its filtered output value. As the filter progresses in the cross-range direction, it is permitted to snake to the right or left (in the range dimension) in order to follow the best path.

A simple example of the filtering process is presented in Figure 2.3. The horizontal axis is the cross-range index and the vertical axis is the image pixel intensity. Note that there are two target-size highlights: one sitting in a nominal background and one sitting on a bright background. Neither the forward nor backward filter can follow the short duration target, but one or the other is able to estimate the longer duration background level. Thus, effective normalization is achieved by choosing the delayed filter output from either the forward or backward filter that is nearest the original intensity to normalize each pixel. Note in the lower plot of Figure 2.3 that the signal-to-background ratios of the targets reflect their respective local background levels in the original image.

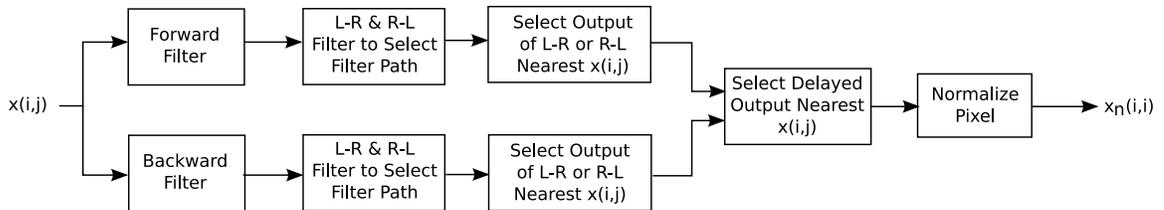


**Figure 2.2:** Illustration of SFBF Path Selection.



**Figure 2.3:** Illustration of the Normalization Method.

The normalizing algorithm follows [1] the block diagram presented in Figure 2.4 where the input image is denoted by  $x(i, j)$  and the normalized image as  $x_n(i, j)$ . For ease of reading the terminology in the algorithm, “bottom-to-top”, “top-to-bottom”, “right-to-left”, and “left-to-right” will be used to describe the path direction relative to an image oriented such that  $x(1, 1)$  is the bottom left corner of the image and  $x(N, M)$  is the top right corner, where the image is of dimension  $N \times M$ . First the forward filter proceeds through the image from bottom to top and the backward filter from top to bottom. For both the forward and backward filters the images are filtered from the left-to-right and right-to-left to select the next pixel in the path of the forward or backward filter.



**Figure 2.4:** Flow Chart of SFBF Normalization Method.

First, the forward filter proceeds through the image from bottom-to-top. The filter is first initialized by finding the mean of the pixels in the current row from the 1st pixel to the  $M_a^{th}$  pixel. The parameter  $M_a$  is defined as  $M_a = \text{Int}[D_a/d_y + 0.5]$ , where  $\text{Int}[\alpha]$  implies the integer part of  $\alpha$  and  $D_a$  is the cross-range distance for average computation and is typically five times the filter correlation distance,  $y_f$  which will be defined later in this section. Therefore, the initial values of the filter are  $y(i, 1) = \frac{1}{M_a} \sum_{j=1}^{M_a} x(i, j)$ ,  $y(i, 2) = y(i, 1)$ ,  $k(i, 1) = i$ ,  $k(i, 2) = i$  for  $i = 1 : N$  and  $k(i, j)$  is the memory for storing the filter path direction.

The filter then proceeds through the image by following the filter recursions as in Tables 2.1 and 2.2. These equations are evaluated for rows  $j = 3 : M$  of the image.

	For Rows $j = 3 : M$
	First Column $i = 1$
	$y(1, j) = c_1 y(1, j - 1) + c_2 y(k(1, j - 1), j - 2) + d_0 x(1, j)$ $u(1) = 1$
	Columns $i = 2 : N$
$l_1$	$y_1(i, j) = c_1 y(i - 1, j - 1) + c_2 y(k(i - 1, j - 1), j - 2) + d_0 x(i, j)$
$l_2$	$y_2(i, j) = c_1 y(i, j - 1) + c_2 y(k(i, j - 1), j - 2) + d_0 x(i, j)$
$l_3$	$y_3(i, j) = c_1 y(u(i - 1), j - 1) + c_2 y(k(u(i - 1), j - 1), j - 2) + d_0 x(i, j)$
	$u(i) = \begin{cases} i - 1 & \text{if pick } l_1 \\ i & \text{if pick } l_2 \\ u(i - 1) & \text{if pick } l_3 \end{cases}$
	$y_{LR}(i, j) = y_k(i, j)$ where $k = 1, 2, \text{ or } 3$

**Table 2.1:** Forward Left to Right Filter Equations in Order of Execution.

The left-to-right (LR) recursions in Table 2.1 find the filter output from the bottom left ( $l_1$ ), bottom ( $l_2$ ), and left ( $l_3$ ) neighboring pixels and choose the one that is nearest the current pixel  $x(i, j)$ . From this choice the filter output is stored to  $y_{LR}(i, j)$  and the path of the output is stored to  $u(i)$ . The right-to-left (RL) recursions in Table 2.2 find the filter output from the bottom right ( $r_1$ ), bottom ( $r_2$ ), and right ( $r_3$ ) neighboring pixels and choose the one that is nearest the current pixel  $x(i, j)$ . From this choice the filter output is stored to  $y_{RL}(i, j)$  and the path of the output is stored to  $v(i)$ .

	For Rows $j = 3 : M$
	Last Column $i = N$
	$y(N, j) = c_1y(N, j - 1) + c_2y(k(N, j - 1), j - 2) + d_0x(N, j)$ $v(N) = N$
	Columns $i = N - 1 : 1$
$r_1$	$y_1(i, j) = c_1y(i + 1, j - 1) + c_2y(k(i + 1, j - 1), j - 2) + d_0x(i, j)$
$r_2$	$y_2(i, j) = c_1y(i, j - 1) + c_2y(k(i, j - 1), j - 2) + d_0x(i, j)$
$r_3$	$y_3(i, j) = c_1y(v(i + 1), j - 1) + c_2y(k(v(i + 1), j - 1), j - 2) + d_0x(i, j)$
	$v(i) = \begin{cases} i + 1 & \text{if pick } r_1 \\ i & \text{if pick } r_2 \\ v(i + 1) & \text{if pick } r_3 \end{cases}$
	$y_{RL}(i, j) = y_k(i, j)$ where $k = 1, 2,$ or $3$

**Table 2.2:** Forward Right to Left Filter Equations in Order of Execution.

The output of either  $y_{LR}(i, j)$  or  $y_{RL}(i, j)$  that is nearest  $x(i, j)$  is chosen and stored as the output of the forward filter as  $y_F(i, j)$  and the path of the output is stored to  $k(i, j) = u(i)$  if the L-R recursion is chosen or  $k(i, j) = v(i)$  if the R-L recursion is chosen.

Second, the backward filter proceeds through the image analogous to that of the forward filter explained above, but the image  $x(i, j)$  is processed from top-to-bottom. As before, the filter outputs are generated for the L-R and R-L recursions and the output of either L-R filter or R-L filter which is nearest  $x(i, j)$  is stored as the output of the backward filter as well as the path to the selected output. The backward recursions starts from the last row and propagates in the reverse direction of the forward recursions. An example of this process is presented in Figure 2.2 where the chosen paths for the forward and backward filter recursions are represented by the red arrows and those that are not paths chosen are represented by the green arrows.

To normalize the pixel, a delay of the filter output  $J_d$  is chosen for both the forward and backward filters. The delay is chosen by  $J_d = \text{Int}[T_s/d_y + 0.5]$ , where  $T_s$  and  $d_y$  will be defined later in the section. Using the delay, the delayed outputs of both the forward ( $y_F(i, j)$ ) and backward ( $y_B(i, j)$ ) filters are found with the corresponding

range index. The delayed output from either the forward filter or the backward filter that is nearest  $x(i, j)$  is chosen to normalize  $x(i, j)$ . The normalized pixel is defined as  $x_n(i, j) = x(i, j)/y(n, j - J_d)$ , where  $y(n, j - J_d)$  is the delayed output of the chosen filter at the corresponding range index  $n$  and cross-range index  $j - J_d$ . The range index  $n$  is found by searching  $k(i, j)$  recursively to determine the range index for the pixel on the forward path or the backward path to  $(i, j)$  that has cross-range index equal to  $j - J_d$ . If the chosen output of the filter is equal to zero then  $x_n(i, j) = 1$ . This is to avoid any numerical instability in the normalization algorithm.

The filter coefficients  $c_1, c_2$  and  $d_0$  in the equations in Tables 2.1 and 2.2 are set as  $c_1 = 2a$ ,  $c_2 = -a^2$ , and  $d_0 = (1 - a)^2$  where  $a = \exp(-d_y/y_f)$ . That is, the value  $a$  is dependent on  $d_y$  and  $y_f$  which are the cross-range resolution of the image and the filter correlation distance, respectively. The parameter  $y_f$  is determined such that objects of size  $T_s$  and smaller are preserved while background patches of size  $B_s$  and larger are normalized out. This is achieved by determining  $y_f$  such that  $G_1(y_f) = G_2(y_f)$ , where  $G_1(y)$  is the figure of merit for target distortion and is defined as,

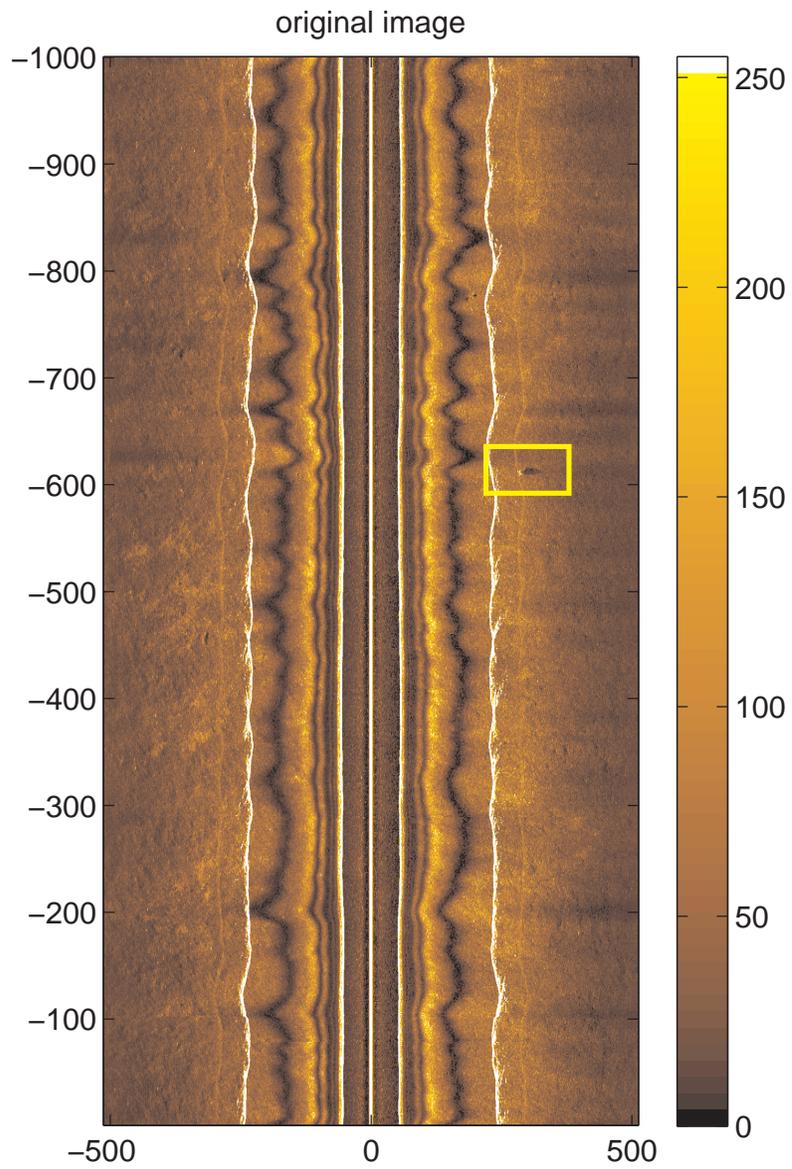
$$G_1(y) = 1 - \left(1 - \frac{0.5T_s}{y}\right) \exp\left(-\frac{0.5T_s}{y}\right) \quad (2.1)$$

and  $G_2(y)$  is defined as the figure of merit for background normalization and is defined as,

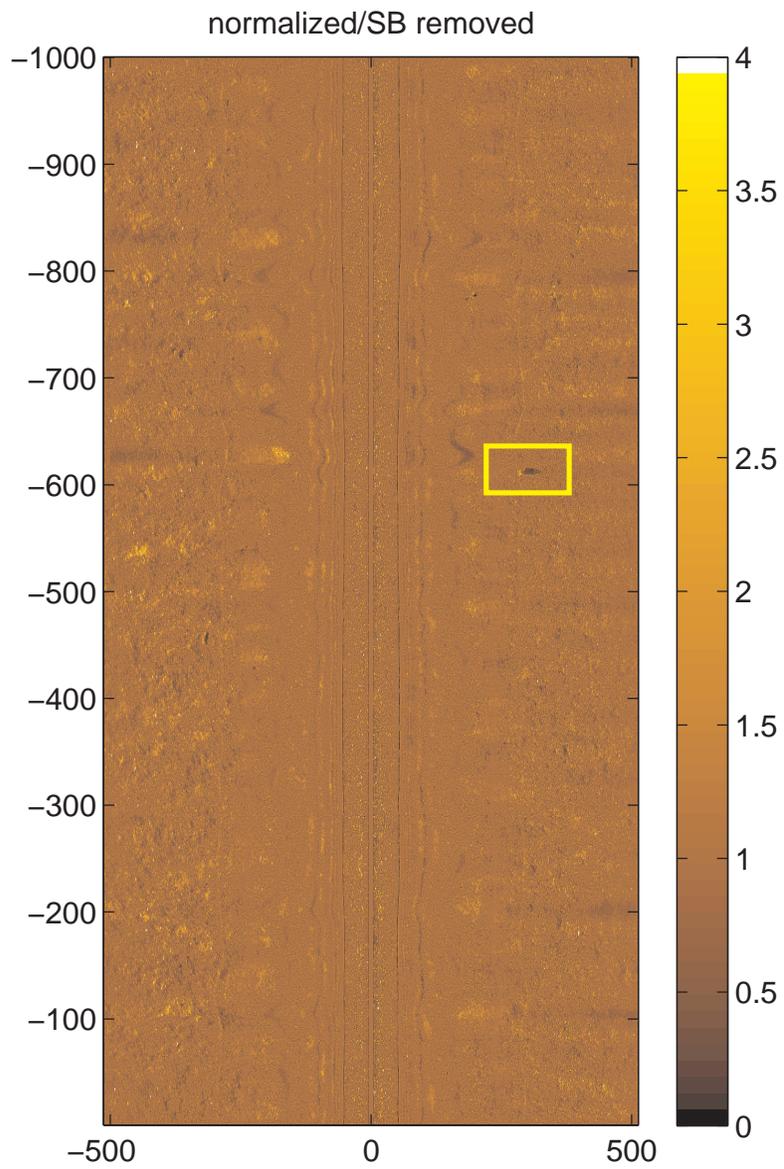
$$G_2(y) = 1 - \left(\frac{0.5B_s - T_s}{y}\right) \exp\left(-\frac{0.5B_s - T_s}{y}\right) \quad (2.2)$$

Small values of  $G$ 's, relative to 1, imply low target distortion and better background normalization. From (2.1) and (2.2) it can be noted that as the difference  $0.5B_s - T_s$  increases, the values of the  $G$ 's decrease. Therefore, if the minimum background patch size, which one desires to normalize out, is sufficiently larger than the maximum target size, then there will be little target distortion and very good background normalization.

Figure 2.5 shows an example of an original sonar image with the target circled on



**Figure 2.5:** Original Side-Scan Sonar Image.



**Figure 2.6:** SFBBF-Normalized Side-Scan Sonar Image.

the image in smooth bottom conditions. The main strip in the center of the image is the area under the sonar and hence not actually part of the image and is not useful. Moving from the L-R of the center of the image the rippled region corresponds to a sharp bottom return near the location of the sonar. The very bright white lines past the rippled region correspond to surface returns of the ocean surface. The rest of the image contains returns from the ocean floor and clutter. As stated previously the purpose of the normalization is to reduce the clutter effects that are present in the sonar image and to enhance the target signature. Figure 2.6 shows the SFBF-normalized image. As can be seen, after the normalization process there is a significant reduction in the amount of clutter as well as an evening out of the background pixel intensity values. Besides the reduction in clutter, the anomalies from the sonar track, surface return, and sharp bottom returns are significantly reduced and partly removed from the image. The target signature is also enhanced with the shadow and highlight being more prominent in the SFBF normalized image.

### 2.3.2 Image Partitioning and CCA-Channelization

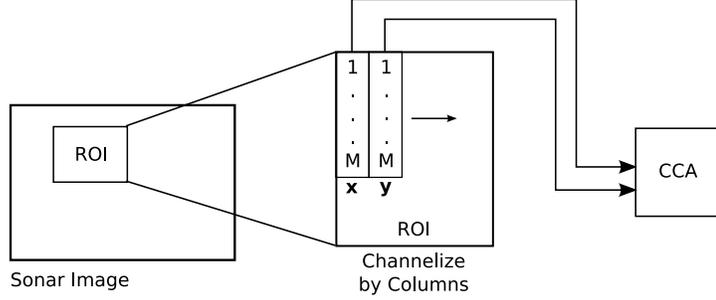
After the normalization process, the first  $N$  columns of pixels are ignored. This region corresponds to the sonar altitude as it travels through the water column and is chosen to be  $1/10^{th}$  of the maximum range of the sonar. The reason behind this is to remove the part of the image that corresponds to the sonar track which carries no information about the environment. Next, the image is partitioned into overlapping ROI's of size  $M \times N$  pixels. The partitioning adopts an overlap along the horizontal and vertical directions of 50% in order to ensure that a target will be covered by more than one ROI in case of splitting the target between ROI's. The size of the ROI is determined experimentally for each data set and is based on the average target size that is encountered in the data set. For the Sonar8 data set, the ROI size is  $46 \times 180$  while for the multi-platform data set the high frequency and broadband

sonar imagery were chosen to be  $72 \times 112$  pixels and  $24 \times 224$ , respectively.

Each ROI is then channelized in order to prepare the data for CCA. The channeling is designed to produce measurements that should be coherent, by virtue of their dependence on a common generation mechanism. In other words, the idea behind this channelization is to look for common coherent attributes that can be used to relate one channel to the other for optimum detection. For the single sensor application, the channels are chosen to be consecutive columns in an ROI where the  $x$  channel is the first column in the ROI and the next column is the  $y$  channel and is shown in Figure 2.7. For the dual disparate sensor application, a pair-wise rectangular vectorization makes more sense due to the two-sensor nature of the problem and the fact that coherence is to be determined between the two sensors data. This process is presented in Figure 2.8 where is shown how each ROI is channelized by a rectangular blocking scheme. The block in the first ROI forms a realization of the  $x$  channel and the block in the co-registered ROI forms the corresponding realization of the  $y$  channel

Once the two channels in CCA are identified, the question is how to vector the channels so that coherence analysis using CCA may be computed for the channel *vectors*. One of the main decisions is how large the dimension of the vectors should be. This is important due to the fact that typically large channel vectors produce a small number of significant canonical correlations, so that a processing gain of  $m/r$  is achieved, where  $m$  is the dimension of the channel vector and  $r$  is the number of significant canonical correlations. But if the dimension of the channel vectors is chosen too large, the coherence can be destroyed. The adequate choice of vectoring would seem to offer the best trade-off between coherent and non-coherent processing. Moreover, the choice of vectoring will be different for different applications of the coherence-based detector.

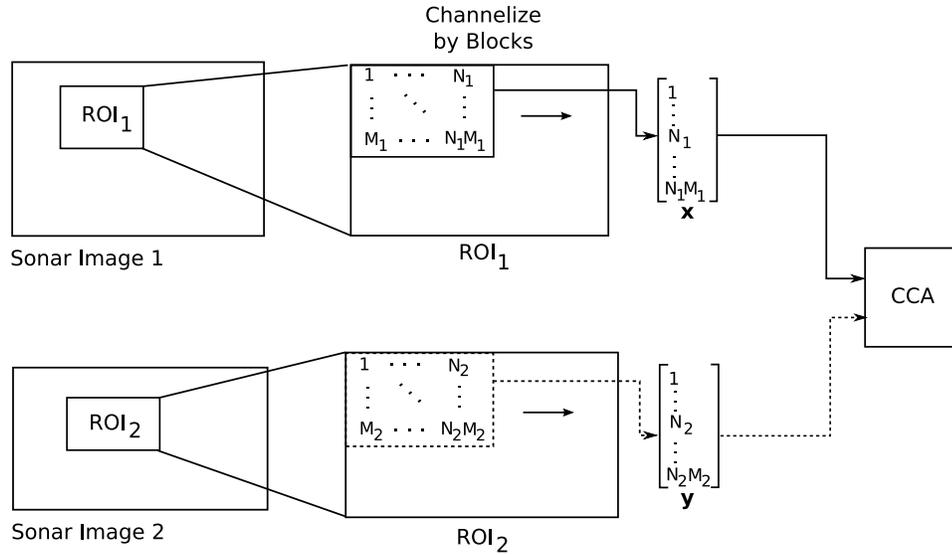
Specifically, for the single sensor application, a column-wise vectorization makes



**Figure 2.7:** Channelization for CCA for the Single Sensor Case.

the most sense. This method is presented in Figure 2.7 where it is shown how each individual ROI is vectorized in a column-wise fashion. The dimension for both the  $\mathbf{x}$  and  $\mathbf{y}$  channels is  $M$  pixels ( $M = 30$  for the Sonar8 data set). The ROI is vectorized such that the  $M$  pixels of the first column form the  $\mathbf{x}$ -channel and the  $M$  pixels in the adjacent column form the  $\mathbf{y}$ -channel vector. This process is continued moving in the horizontal direction across the ROI. On the next pass through the ROI, the channels are given a 50% overlap in the vertical direction to ensure complete coverage of the target in the ROI. Clearly, for background ROI's high level of coherence among consecutive columns (channels) does not exist and by using columns, specific coherence is found between specific parts of the target structure.

For the dual disparate sensor application, dimension of the blocks, which form the channels, for the first and second co-registered sonar images, are  $M_1 \times N_1$  and  $M_2 \times N_2$  pixels, respectively. The channel vectors are formed by row-wise arrangement of the block into  $M_1 N_1 \times 1$  and  $M_2 N_2 \times 1$  vector as shown in Figure 2.8. This process is repeated across both ROI's moving in the horizontal direction to vectorize the ROI. Clearly, for a pair of background ROI's a high level of coherence between two ROI's is not expected to exist. For the multi-platform data set, the block sizes for the high frequency and broadband sonar images were chosen to be  $6 \times 4$  and  $2 \times 8$  pixels, respectively. These sizes were chosen based upon the ROI size and to provide enough samples of the  $\mathbf{x}$  and  $\mathbf{y}$  channels to avoid data poverty issues. The ROI



**Figure 2.8:** Channelization for CCA for the Dual Disparate Sensor Case.

sizes are decided based upon the resolution of the sonar and the average target size encountered in the data set.

Finally, once the ROI's are channelled and vectored an averaging scheme must be adopted to compute the sample composite covariance of the two channel data. Ideally, we would like to identify multiple, independent copies of each channel vector in the ROI. However, this is idealized, physical properties of the target, medium would conspire to defeat it. What is required here is a methodology for choosing the averaging such that we stay as close to the independent assumption as possible.

The averaging for both of the applications is similar. For the single sensor case, the columns are moved in the horizontal direction across the ROI. On the next pass through the same ROI, the channel vectors are shifted down vertically by 50% to ensure complete coverage of the target within the ROI (see Figure 2.7). For the dual disparate case, the blocks (or channel vectors) in the two co-registered ROI's are moved in the horizontal direction across the ROI and the process is repeated until the entire ROI has been covered (see Figure 2.8). After the data matrices for the two channels have been formed, they are used to form the sample covariance matrices

that make up the coherence matrix in the CCA process.

## 2.4 Conclusion

In this chapter, the two sonar data sets that are used in this thesis were reviewed. Both data sets consist of side-scan sonar imagery that cover various target fields containing varying degrees of background clutter and bottom difficulty. The Sonar8 data set consists of images captured by a high resolution side looking sonar. The Sonar8 image with pixels represent the envelope data, or simply the magnitude of the complex output of the beamformer. This data set provides various levels of background clutter and bottom difficulty with some targets being obscured by very dense and textured clutter. The multi-platform sonar data set provides a disparate sensory setup in which a high frequency sonar and three broadband sonars that are co-registered over the same region of the target field are used. The use of multiple disparate sonars allows one to exploit the good characteristics of both sonar types at the same time during the detection and classification. A high resolution, high frequency sonar provides very good target definition, but at the same time generates a large number of contacts. Broadband sonars reduce the clutter, while sacrificing target definition and resolution. Therefore, by using both types of sonar systems co-registered over the same region one can provide potentially much better detection results compared to those of the single sonar (Sonar8) cases.

To help enhance detection of targets in varying background conditions the SFBF normalization algorithm is applied to each image prior to detection. The purpose of this normalization is to help emphasize the target's highlight and shadow signatures to better distinguish them from the bottom clutter present in the image. This is done by reducing the variability of the local mean throughout the image. The local mean is then used to normalize each pixel so that the highlight and shadow of the target can be more easily identified and the artifacts in the sonar image are reduced.

After normalization the sonar images are partitioned into ROI's with a 50% overlap. Each ROI is then channelized such that the channels are identified for the CCA processing. For the single sensor application the channels are consecutive columns in the ROI and for the dual disparate sensor case the channels are the pair of blocks from the co-registered ROI's. Following the channelization the ROI's are vectorized such that channel dimension is chosen and the samples of the channel are extracted and reshaped into a column vector for the realizations of the data channels for CCA. The vectors are then averaged to compute the sample composite covariance of the two channel data for CCA. In the next chapter, Gauss-Gauss detection [29], [35] in the CCA framework is reviewed.

# CHAPTER 3

## BINARY HYPOTHESIS DETECTION METHODS

### 3.1 Introduction

The standard detection problem involves evaluating a set of hypotheses, to determine which is the most likely with the most basic being the binary hypothesis test. The binary hypothesis test is described as a decision between either the “true” hypothesis  $H_1$ , or the alternative “null” hypothesis  $H_0$ . This is typically viewed as a signal detection problem where  $H_1$  is considered as the signal plus noise case and  $H_0$  is considered as the noise only case. The most common framework for hypothesis testing is the Bayesian framework [47] - [49] where the problem is formulated by minimizing the Bayes risk. When the costs and *a priori* probability densities are available, Bayesian detection is optimal, but in general it is difficult to generate these. The Neyman-Pearson criterion [47] - [49] offers an alternative to the standard Bayesian framework by formulating the hypothesis test as a constrained optimization problem where the false alarm probability is constrained and the probability of detection is maximized. The solution to this optimization problem is the celebrated Neyman-Pearson lemma. Solving the optimization problem leads to a likelihood ratio test and threshold that does not depend on the *a priori* probabilities of the problem. If the assumption is made that under both hypotheses the data is Gaussian distributed the likelihood ratio test then becomes the standard Gauss-Gauss detector [35].

In developing a detection framework for sonar imagery the *a priori* probability

densities are difficult to determine due to variations in operating and environmental conditions. Due to the difficulty of the problem, choosing the right detection framework and correct coordinate system is crucial. Canonical correlation analysis (CCA) [27] provides an excellent framework for coherence-based detection. The CCA method determines linear dependence [27] (or coherence) between two data channels. This method not only determines the amount of dependence (or independence) between two data channels (e.g. columns in a ROI within a sonar image or blocks in a pair of ROI's of two sonar images) but also extracts, via the canonical coordinates, a subset of the most coherent features for classification purposes. Canonical coordinate decomposition allows us to quantify the changes between the returns from the bottom and when target activities are present and at the same time extract useful features for target classification without the need to perform separate detection and anomaly feature extraction. Moreover, the standard Gauss-Gauss detector can alternatively be cast in the CCA framework [29] allowing for a detector that uses the coherence information between two data channels. This detector lends itself nicely to the detection of underwater objects from sonar imagery [14] - [15], as the targets tend to be more coherent than the environment in which they are found. Specifically, when a target is present in an ROI (or two co-registered ROI's) the coherence between consecutive columns (channels) of pixels are more coherent than when background is present only. The canonical correlations provide measure of the coherence (or incoherence) to determine if a target (or background clutter) is present in the processed ROI.

The outline of this chapter is as follows. Section 3.2 opens the chapter with a brief review of Bayesian detection framework. Section 3.2.1 presents the Neyman-Pearson criterion and develops the optimum Gauss-Gauss detector. Section 3.3 presents a review of CCA-based Gauss-Gauss detection [29] and its benefits. Finally, conclusions are made in Section 3.4.

## 3.2 Bayesian Detection Framework

Consider the classical detection problem of choosing between two hypotheses [47] where each hypothesis relates to a point in the  $N$ -dimensional observation space. In other words, if  $\mathbf{x} = [x_1, x_2, \dots, x_N]^H$ , is an observation vector in this space we would like to test between  $H_1$  hypothesis (true) and  $H_0$  hypothesis (null) for this observation vector. Clearly, each time we conduct the test there are four possible outcomes. Those are: (a)  $H_0$  is true and we choose  $H_0$ , (b)  $H_0$  is true and we choose  $H_1$ , (c)  $H_1$  is true and we choose  $H_1$ , and (d)  $H_1$  is true but we choose  $H_0$ . The first and third outcomes lead to correct decisions while the second and fourth outcomes lead to erroneous decisions. The Bayes test is based on two assumptions. First, the two hypotheses,  $H_0$  and  $H_1$ , correspond to two possible prior probabilities,  $P_0$  and  $P_1$ , respectively. These probabilities represent the prior observer's information about the hypotheses before the detection is conducted. The second assumption is that there is a cost associated with each of the four courses of action described above. These costs will be denoted by,  $C_{00}, C_{10}, C_{11}$ , and  $C_{01}$ , for outcomes 1-4, respectively. It will be assumed that the cost of a wrong decision is higher than the cost of a correct decision, i.e.  $C_{10} > C_{00}$  and  $C_{01} > C_{11}$ . The goal of the Bayes test is to design a decision rule so that on the average the cost of a decision will be as small as possible, which subsequently leads to the smallest Bayesian risk when making the decision. If we denote the expected value of the cost as the risk  $\mathcal{R}$ , we can then write  $\mathcal{R}$  as [47],

$$\begin{aligned}\mathcal{R} &= C_{00}P_0P(H_0|H_0) \\ &+ C_{10}P_0P(H_1|H_0) \\ &+ C_{11}P_1P(H_1|H_1) \\ &+ C_{01}P_1P(H_0|H_1)\end{aligned}\tag{3.1}$$

where  $P(H_j|H_i)$   $i, j \in [0, 1]$  is the probability that we choose  $H_j$  given that the true hypothesis is  $H_i$ .

Since the decision rule is binary, i.e. there are only two possibilities, either  $H_0$  and  $H_1$ , we can view the rules as a division in the observation space into two parts  $A_0$  and  $A_1$ . In other words, if the observation is found in the region  $A_0$  the hypothesis  $H_0$  is declared true and if the observation is found in the region  $A_1$  the hypothesis  $H_1$  is declared true. By viewing the problem in this manner we can now express the risk in terms of the decision regions and probabilities as,

$$\begin{aligned}
\mathcal{R} &= C_{00}P_0 \int_{A_0} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} \\
&+ C_{10}P_0 \int_{A_1} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} \\
&+ C_{11}P_1 \int_{A_1} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} \\
&+ C_{01}P_1 \int_{A_0} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x}. \tag{3.2}
\end{aligned}$$

To find the decision rule, the decision regions are determined such that the risk in (3.2) is minimized. Because each element of  $\mathbf{x}$  must be assigned to either the  $A_0$  or  $A_1$  in the observation space  $A$ , we can say that  $A = A_0 \cup A_1$  and  $A_0 \cap A_1 = \emptyset$ . Now, (3.2) can be rewritten as [47]

$$\begin{aligned}
\mathcal{R} &= P_0C_{00} \int_{A_0} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} + P_0C_{10} \int_{A-A_0} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} \\
&+ P_1C_{01} \int_{A_0} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} + P_1C_{11} \int_{A-A_0} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x}. \tag{3.3}
\end{aligned}$$

We can separate the integrals and rewrite (3.3) as,

$$\begin{aligned}
\mathcal{R} &= P_0C_{00} \int_{A_0} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} + P_0C_{10} \int_A p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} \\
&- P_0C_{10} \int_{A_0} p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} + P_1C_{01} \int_{A_0} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} \\
&+ P_1C_{11} \int_A p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} - P_1C_{11} \int_{A_0} p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} \tag{3.4}
\end{aligned}$$

If we use  $\int_A p_{\mathbf{X}|H_0}(\mathbf{x}|H_0) d\mathbf{x} = \int_A p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) d\mathbf{x} = 1$ , then (3.4) can be reduced to,

$$\begin{aligned}
\mathcal{R} &= P_0C_{10} + P_1C_{11} \\
&+ \int_{A_0} [P_1(C_{01} - C_{11})p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) - P_0(C_{10} - C_{00})p_{\mathbf{X}|H_0}(\mathbf{x}|H_0)] d\mathbf{x} \tag{3.5}
\end{aligned}$$

The first two terms in (3.5) represent the fixed cost and the integral represents the cost controlled by the points in the observation space,  $A$  that are assigned to  $A_0$ . Basically, the points in  $A$  for which the first term in the integral is larger than the second term are assigned to  $A_1$ , whereas the points in which the second term is larger than the first term are assigned to  $A_0$ . Any points in which the terms are equal have no effect on the cost and can be arbitrarily assigned to any region (we assume that these points are assigned to  $A_1$ ). We can, therefore, define the decision region in the observation space by

$$P_1(C_{01} - C_{11})p_{\mathbf{X}|H_1}(\mathbf{x}|H_1) \geq P_0(C_{10} - C_{00})p_{\mathbf{X}|H_0}(\mathbf{x}|H_0). \quad (3.6)$$

which can alternatively be rewritten as

$$\frac{p_{\mathbf{X}|H_1}(\mathbf{x}|H_1)}{p_{\mathbf{X}|H_0}(\mathbf{x}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}. \quad (3.7)$$

The quantity on the left is called the *likelihood ratio* and will be denoted throughout the rest of this thesis by

$$l(\mathbf{x}) \triangleq \frac{p_{\mathbf{X}|H_1}(\mathbf{x}|H_1)}{p_{\mathbf{X}|H_0}(\mathbf{x}|H_0)}. \quad (3.8)$$

The relationship on the right is the threshold of the test and will be denoted by  $\eta$ . Thus, Bayes criterion simply leads to a likelihood ratio test,

$$l(\mathbf{x}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad (3.9)$$

### 3.2.1 Neyman-Pearson Detection

In many practical detection situations it is difficult to assign realistic costs and prior probabilities. This is especially apparent for sonar imagery target detection due to the variations in target compositions and background clutter. To overcome this, a simple alternative is to use probabilities,  $P_{fa}$ , i.e. false alarm probability and,  $P_d$ , which is the probability of detection. This leads to the Neyman-Pearson criterion [47], [50] which designs a test to maximize  $P_d$  while making  $P_{fa}$  as small as possible. The

criterion constrains  $P_{fa} = \alpha' \leq \alpha$  and designs a test that maximizes the probability of detection under this constraint [47]. The solution to this constrained optimization problem is found by using Lagrange multipliers, i.e., we can construct a function  $F$ ,

$$F = P_d + \lambda[P_{fa} - \alpha']. \quad (3.10)$$

Here by maximizing  $F$  we will maximize  $p_D$ . For ease in the rest of the derivation we will convert this to a minimization problem where the objective is to minimize the probability of a miss-detection,  $P_m$  which gives  $F = P_m + \lambda[P_{fa} - \alpha']$ . We can then rewrite  $F$  in terms of the conditional probabilities as [47]

$$F = \int_{A_0} p_{\mathbf{x}|H_1}(\mathbf{x}|H_1) d\mathbf{x} + \lambda \left[ \int_{A_1} p_{\mathbf{x}|H_0}(\mathbf{x}|H_0) d\mathbf{x} - \alpha' \right]. \quad (3.11)$$

Clearly (3.11) is minimized when  $P_{fa} = \alpha'$ , thus minimizing  $P_m$  for the non-trivial case when  $\lambda \neq 0$ . Using  $A_0 = A - A_1$ , we can then rewrite  $F$  as

$$F = \lambda(1 - \alpha') + \int_{A_0} [p_{\mathbf{x}|H_1}(\mathbf{x}|H_1) - \lambda p_{\mathbf{x}|H_0}(\mathbf{x}|H_0)] d\mathbf{x} \quad (3.12)$$

For any positive value of  $\lambda$  a likelihood ratio test will minimize  $F$  [47]. This follows directly because by minimizing  $F$  we assign a point in  $\mathbf{x}$  to  $A_1$  when the integral is positive or assign the point to  $A_0$  when the integral is negative. Therefore, we can write

$$l(\mathbf{x}) = \frac{p_{\mathbf{x}|H_1}(\mathbf{x}|H_1)}{p_{\mathbf{x}|H_0}(\mathbf{x}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda. \quad (3.13)$$

Although this is similar to (3.9), the threshold in (3.13) does not rely on the *prior* probabilities and the assigned costs.

### 3.2.2 Gauss-Gauss Signal Detection

If we now view this detection problem in terms of the signal plus noise model [35] - [29], the decision between two hypotheses is now either noise only ( $H_0$ ) or signal plus noise ( $H_1$ ). Assume we have an observation  $\mathbf{x} \in \mathbb{R}^N$ , which is a normal random vector with zero mean and covariance matrix  $R$ . We wish to test the hypothesis

$H_0 : R = R_0$ , i.e. noise alone versus  $H_1 : R = R_1$ , i.e signal plus noise, where  $R_1 = R_0 + R_s$ ,  $R_0$  is the covariance matrix of the noise alone, and  $R_s$  is the covariance matrix of the signal. Note that it is assumed that noise and signal are uncorrelated.

Since we have assumed that  $\mathbf{x}$  is Gaussian with zero mean and covariance matrix  $R$ , the probability density function for a given hypothesis  $H_i$ ,  $i \in [0, 1]$  and a given observation  $\mathbf{x}$  is given by

$$p_{\mathbf{x}|H_i}(\mathbf{x}|H_i) = (2\pi)^{-\frac{N}{2}} |R_i|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{x}^H (R_i^{-1})\mathbf{x}}. \quad (3.14)$$

Using (3.8) and taking the natural log, the log-likelihood of  $\mathbf{x}$  becomes

$$\begin{aligned} l(\mathbf{x}) &= \ln \left( \frac{(2\pi)^{-\frac{N}{2}} |R_1|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{x}^H R_1^{-1}\mathbf{x}}}{(2\pi)^{-\frac{N}{2}} |R_0|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{x}^H R_0^{-1}\mathbf{x}}} \right) \\ &= \ln \left( \frac{|R_1|^{\frac{1}{2}}}{|R_0|^{\frac{1}{2}}} e^{\frac{1}{2}\mathbf{x}^H (R_0^{-1} - R_1^{-1})\mathbf{x}} \right) \\ &= \frac{1}{2} \ln |R_1| - \frac{1}{2} \ln |R_0| + \frac{1}{2} \mathbf{x}^H (R_0^{-1} - R_1^{-1}) \mathbf{x} \end{aligned} \quad (3.15)$$

Disregarding the constants that are not observation dependent, the log-likelihood can be simply written as

$$l(\mathbf{x}) = \mathbf{x}^H (R_0^{-1} - R_1^{-1}) \mathbf{x}. \quad (3.16)$$

which is the likelihood-ratio for the Gauss-Gauss detector [35].

The log-likelihood ratio in (3.16) minimizes the risk involved in deciding between  $H_0$  and  $H_1$  and can be viewed as,

$$\begin{aligned} \gamma(\mathbf{x}) &= \begin{cases} 1 \sim H_1, & l(\mathbf{x}) > \lambda \\ 0 \sim H_0, & l(\mathbf{x}) \leq \lambda \end{cases} \\ l(\mathbf{x}) &= \mathbf{x}^H \mathbf{Q} \mathbf{x} \text{ with } \mathbf{Q} = R_0^{-1} - R_1^{-1}, \end{aligned}$$

where  $\lambda$  is the detection threshold and  $\mathbf{Q}$  can be written as [35]

$$\mathbf{Q} = R_0^{-H/2} (\mathbf{I} - \mathbf{S}^{-1}) R_0^{-1/2}, \quad (3.17)$$

and  $S$ , which is known as the “signal-to-noise ratio” matrix [35], is defined as

$$S = R_0^{-1/2} R_1 R_0^{-H/2} \quad (3.18)$$

Now, we can rewrite  $l(\mathbf{x})$  as,

$$l(\boldsymbol{\xi}) = \boldsymbol{\xi}^H (I - S^{-1}) \boldsymbol{\xi} \quad (3.19)$$

where,  $\boldsymbol{\xi} = R_0^{-1/2} \mathbf{x}$ .

Clearly, the transformed vector  $\boldsymbol{\xi}$  is also normal with zero mean and covariance matrix  $R = I$  under  $H_0$ , and  $R = S$  under  $H_1$  i.e.  $E_{H_0}[\boldsymbol{\xi}\boldsymbol{\xi}^H] = I$  and  $E_{H_1}[\boldsymbol{\xi}\boldsymbol{\xi}^H] = S$ , where  $E_{H_0}[\cdot]$  is the expectation operation under  $H_0$ , and  $E_{H_1}[\cdot]$  is the expectation operation under  $H_1$ .

The divergence [35] between the two hypotheses, which can be used as a measure of detectability, is

$$\begin{aligned} J &= E_{H_1}[l(\boldsymbol{\xi})] - E_{H_0}[l(\boldsymbol{\xi})] \\ &= \text{tr}(I - S^{-1})E_{H_1}[\boldsymbol{\xi}\boldsymbol{\xi}^H] - \text{tr}(I - S^{-1})E_{H_0}[\boldsymbol{\xi}\boldsymbol{\xi}^H] \\ &= \text{tr}(I - S^{-1})S - \text{tr}(I - S^{-1})I = \text{tr}(S + S^{-1} - 2I), \end{aligned} \quad (3.20)$$

where  $\text{tr}(\cdot)$  denotes the trace operation on a matrix. The matrix  $S$  can be decomposed as

$$S = R_0^{-1/2} R_1 R_0^{-H/2} = U\Lambda U^H, \quad (3.21)$$

where  $\Lambda$  is a diagonal matrix with diagonal elements  $\lambda_i$ , which are the eigenvalues of  $S$ , and  $U$  is an eigenvector matrix containing the corresponding eigenvectors in its column space. This implies that  $(R_0^{-H/2}U, \Lambda)$  solves the following generalized eigenvalue problem [35]

$$R_1(R_0^{-H/2}U) = R_0(R_0^{-H/2}U)\Lambda. \quad (3.22)$$

Using the cyclic property of the trace, the divergence in (3.21) can be rewritten as

$$\begin{aligned} J &= \text{tr}(\Lambda + \Lambda^{-1} - 2\mathbf{I}) \\ &= \sum_{i=1}^N (\lambda_i + \lambda_i^{-1} - 2). \end{aligned} \quad (3.23)$$

Using the eigenvalue decomposition of  $\mathbf{S}$  the log-likelihood ratio may be rewritten as

$$l(\boldsymbol{\xi}) = \boldsymbol{\xi}^H \mathbf{U}(\mathbf{I} - \Lambda^{-1})\mathbf{U}^H \boldsymbol{\xi}. \quad (3.24)$$

Thus, we only need to solve the eigenvalue problem in (3.22) to obtain the eigenvectors,  $\mathbf{u}_i$ , and eigenvalues  $\lambda_i$  for computing the log-likelihood in (3.24) and the J-divergence in (3.23).

A reduced rank version of the log likelihood ratio is [35]

$$l_r(\boldsymbol{\xi}) = \boldsymbol{\xi}^H \mathbf{U}(\mathbf{I}_r - \Lambda_r^{-1})\mathbf{U}^H \boldsymbol{\xi}, \quad (3.25)$$

where  $\mathbf{I}_r$  and  $\Lambda_r$  are the reduced rank versions of  $\mathbf{I}$  and  $\Lambda$  containing  $r$  non-zero diagonal elements, i.e.  $\mathbf{I}_r = \text{diag}[1, \dots, 1, 0, \dots, 0]$  and  $\Lambda_r = \text{diag}[\lambda_1, \dots, \lambda_r, 0, \dots, 0]$ .

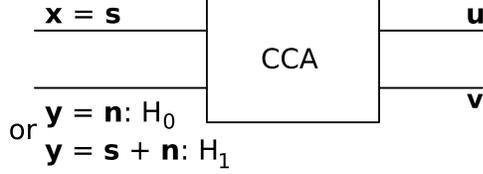
The divergence is then,

$$J = \sum_{i=1}^r (\lambda_i + \lambda_i^{-1} - 2). \quad (3.26)$$

The goal of the reduced rank case is to select the “dominant” eigenvalues, that maximally contributes to the separation between the decision regions in the observation space. The rank- $r$  divergence is identical to the full rank divergence when  $N - r$  of the  $\lambda$ 's are unity. Therefore, by reducing the rank of the detector we increase our processing gain, by reducing the number of multiplications and additions in the log-likelihood, while still maximizing the divergence.

### 3.3 Gauss-Gauss Signal Detection in CCA Framework

The Gauss-Gauss detector in Section 3.2.2 can also be cast in the CCA framework in which the detection test and detection criterion are formed in terms of the canonical



**Figure 3.1:** CCD used as an optimal detector.

coordinates and canonical correlations [29]. For a review of CCA the reader is referred to Appendix A. Canonical correlation analysis offers an ideal framework for coherent-based detection and feature extraction. The extracted canonical correlations for the two data channels provide a coherence (or incoherence) measure that can be used to determine if a target is present (or absent). Therefore, only the dominant correlations need to be retained, which measure the level of coherence, to build a rank- $r$  detector that maximizes the divergence. In the two-channel CCA framework, channel  $\mathbf{x} \in \mathbb{R}^N$  corresponds to signal alone data i.e.  $\mathbf{x} = \mathbf{s}$  with covariance matrix  $R_{xx} = R_s$  whereas channel  $\mathbf{y} \in \mathbb{R}^N$  could be either noise under  $H_0$  (i.e.  $\mathbf{y} = \mathbf{n}$ ) with covariance matrix  $R_{yy} = R_0$  or signal plus noise under  $H_1$  (i.e.  $\mathbf{y} = \mathbf{s} + \mathbf{n}$ ) with covariance matrix  $R_{yy} = R_s + R_0$ . Figure 3.1 shows how the two-channel CCA can be used as an optimal detector. Under hypothesis  $H_1$ , the composite vector  $\mathbf{z} = [\mathbf{x}^H \mathbf{y}^H]^H$  has the composite covariance matrix

$$R_{zz} = E[\mathbf{z}\mathbf{z}^H] = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = \begin{bmatrix} R_s & R_s \\ R_s^H & R_1 = R_s + R_0 \end{bmatrix}. \quad (3.27)$$

In the CCA framework the “signal-to-noise ratio” matrix in (3.18), under hypothesis  $H_1$ , can be rewritten as,

$$S = R_0^{-1/2}(R_s + R_0)R_0^{-H/2} \quad (3.28)$$

and the squared coherence matrix [27],  $CC^H$ , can be expressed as,

$$CC^H = R_s^{H/2}(R_s + R_0)^{-1}R_s^{1/2} \quad (3.29)$$

Therefore, using the log-likelihood (3.19),  $l(\mathbf{y}) = (R_0^{-1/2}\mathbf{y})^H(\mathbf{I} - \mathbf{S}^{-1})(R_0^{-1/2}\mathbf{y})$  inserting for  $\mathbf{S}$  from (3.28), we can rewrite the log likelihood as [29],

$$l(\mathbf{y}) = (R_s^{-1/2}\mathbf{y})^H(R_s^{H/2}R_0^{-1}R_s^{1/2} - R_s^{H/2}(R_s + R_0)^{-1}R_s^{1/2})(R_s^{-1/2}\mathbf{y}). \quad (3.30)$$

Using the fact that  $[(\mathbf{C}\mathbf{C}^H)^{-1} - \mathbf{I}]^{-1} = R_s^{H/2}R_0^{-1}R_s^{1/2}$ , the log-likelihood may be rewritten as

$$l(\mathbf{y}) = (R_s^{-1/2}\mathbf{y})^H([\mathbf{C}\mathbf{C}^H]^{-1} - \mathbf{I})^{-1} - \mathbf{C}\mathbf{C}^H)(R_s^{-1/2}\mathbf{y}). \quad (3.31)$$

Using the SVD of the coherence matrix  $\mathbf{C} = R_s^{H/2}R_1^{-H/2} = \mathbf{F}\mathbf{K}\mathbf{G}^H$  and  $\mathbf{F}^H R_s^{-1/2} = \mathbf{K}^{-1}\mathbf{G}^H R_1^{-1/2}$ , we can write the log-likelihood as

$$l(\mathbf{y}) = (\mathbf{G}^H R_1^{-1/2}\mathbf{y})^H([\mathbf{I} - \mathbf{K}^2]^{-1} - \mathbf{I})(\mathbf{G}^H R_1^{-1/2}\mathbf{y}). \quad (3.32)$$

Alternatively, we have

$$l(\mathbf{v}) = \mathbf{v}^H([\mathbf{I} - \mathbf{K}^2]^{-1} - \mathbf{I})\mathbf{v} \quad (3.33)$$

Thus, the log-likelihood ratio is expressed in terms of the canonical coordinate vector  $\mathbf{v} = \mathbf{G}^H R_1^{-1/2}\mathbf{y}$  for the  $\mathbf{y}$  channel and the canonical correlation matrix  $\mathbf{K}$ . More specifically,  $l(\mathbf{v})$  is the weighted sum of the magnitude-squared of the canonical coordinates weighted by canonical correlation-dependent weights, i.e.

$$l(\mathbf{v}) = \sum_{i=1}^N |v_i|^2 \left( \frac{k_i^2}{1 - k_i^2} \right), \quad (3.34)$$

where  $v_i = \mathbf{g}_i^H R_1^{-1/2}\mathbf{y}$  and  $\mathbf{g}_i$  is the  $i^{\text{th}}$  column of the matrix  $\mathbf{G}$ .

It can also be shown [29] that the  $J$ -divergence between  $H_1$  and  $H_0$  can be expressed solely in terms of the  $\mathbf{K}$  matrix or canonical correlations as

$$J = \text{tr}([\mathbf{I} - \mathbf{K}^2]^{-1} - \mathbf{I} - \mathbf{K}^2) = \sum_{i=1}^N \frac{k_i^4}{1 - k_i^2}. \quad (3.35)$$

The function  $\frac{k_i^4}{1 - k_i^2}$  is non-increasing in the interval  $(0, 1]$ . *Consequently the rank- $r$  detector that maximizes the divergence is the detector that uses the canonical coordinates corresponding to the  $r$ -dominant canonical correlations  $k_1, \dots, k_r$ .* The  $J$ -divergence between the two hypotheses considering  $r$  dominant canonical correlations

is then,

$$J_r = \sum_{i=1}^r \frac{k_i^4}{1 - k_i^2}. \quad (3.36)$$

Thus, for building low-rank detectors, only the dominant canonical coordinates need to be retained in order to find the coherence between the two data channels  $\mathbf{x}$  and  $\mathbf{y}$ .

CCA provides an ideal framework for simultaneous target detection and coherent feature extraction in sonar images. The extracted canonical correlations for either a pair of columns in an ROI or a pair of ROI's provide a coherence measure that can be used to determine if a target is present in the processed ROI(s). Our *detection hypothesis* for target detection in sonar imagery is that presence of objects in a sonar image leads to high level coherence measure comparing to that of the background clutter only. This makes the Gauss-Gauss detector cast in the CCA framework as an ideal tool for detecting objects in sonar imagery. However, due to the nature of the detection problem in sonar imagery the hypotheses will contain two channels. Therefore, in order to use the Gauss-Gauss detector the likelihood ratio and J-divergence will have to be reformulated. This development is presented in Chapter 4.

### 3.4 Conclusion

In this chapter, optimum Bayesian detection and the Neyman-Pearson criterion were reviewed. Bayesian detection is optimal when the decision costs and prior probabilities are known. However, in general it is difficult to generate these values due to lack of *a priori* information about the environment and distribution of true and null hypothesis. The Neyman-Pearson criterion provides a decision rule for testing hypothesis in which the decision costs and prior probabilities are not required. In the Neyman-Pearson criterion, if the problem is viewed as a Gaussian signal and noise detection problem, Gauss-Gauss detector is obtained. In this method the likelihood ratio and J-divergence can be reformulated in terms of the eigenvalues and eigenvectors of the “signal-to-noise ratio” matrix, where rank reduction can be performed

while still maximizing the divergence.

The Gauss-Gauss detector was then cast in the CCA framework. This framework offers ease of implementation and suitability for building a coherence-based detector. The latter is extremely desirable for the detection of targets in sonar imagery as our detection hypothesis is that when a target is present in an ROI, consecutive columns for the single sonar or pair of blocks in two ROI's for dual disparate sonar the coherence is higher than when background only is present. The extracted canonical correlations, which measure the coherence between those two data channels can also be used for target classification.

In Chapter 4, the use of the Gauss-Gauss decision rule is extended to the composite channel hypothesis testing. In other words, when the target detection involves more than one sensors the standard hypothesis testing does not directly apply. This occurs for the single sensor case as the two channels for the detector are two consecutive columns in a ROI. This also occurs for the dual disparate sensor case as a pair of blocks from a pair of co-registered ROI's are the channels for the detector. Since the structure of the two hypotheses have changed new formulations for the log-likelihood ratio and J-divergence need to be developed for the Gauss-Gauss detector to account for this composite two-channel nature of this problem.

# CHAPTER 4

## DETECTION OF TARGETS IN SONAR IMAGERY

### 4.1 Introduction

Detection and classification of underwater objects in sonar imagery is a complicated problem due to various factors such as variations in operating and environmental conditions, presence of spatially varying clutter, variations in target shapes, compositions and orientation. Moreover, bottom features such as coral reefs, sand formations, and the vegetation may totally obscure a target object.

In this chapter a new coherent-based detection method for sonar imagery is developed using CCA framework in Gauss-Gauss detection. Canonical coordinate decomposition allows one to quantify the changes between the returns from the bottom and any target activities in sonar images while at the same time extract useful features without the need to perform separate anomaly feature extraction. These features can be used for simultaneous detection and classification of mine-like and non-mine-like objects.

Current detection methods normally use a single sonar where only one sonar image is used to detect potential targets. Detection based off one image can lead to unacceptable results as the information is limited to the field of view of one sonar. This motivates the use of dual disparate sensors where the sonar platforms could be disparate in either frequency, resolution, location, etc to better capture the target characteristics. By using multiple disparate sonar types allows one to use a high resolution sonar with good target definition and the clutter suppressing abilities of a

low resolution sonar co-registered over the same region. The use of multiple disparate sonar types allows for the processing of all information available at once to make a detection call.

When considering the detection problem in sonar imagery for either the single sonar or the dual disparate sonar cases the development of the CCA-based detector in Section 3.3 does not directly apply. This is due to the fact that for these cases the hypothesis  $H_0$  corresponds to the case when both channels of the detector contain background noise only while  $H_1$  corresponds to the case where both channels contain target (signal) and background noise though the way the channels are defined are different for the single and dual sonar problems. Therefore, new expressions for the log-likelihood and J-divergence need to be developed for this composite channel hypotheses case. However, to reformulate this problem and relate it to the standard case one has to develop the log-likelihood ratio and J-divergence under the new hypotheses. Since in this case the channels are of higher dimensions finding inverses of the covariance matrices is not an efficient approach. In this chapter, we develop a method to find the inverses of the covariance matrices under the  $H_0$  and  $H_1$  hypotheses and relate the composite two-channel hypothesis problem back to the standard one channel problem. Through this relationship we can also show the relationship to the standard CCA-based detector [29] for ease in implementation. We then apply the detector to the single sensor and the dual disparate sensor cases and demonstrate that the use of multiple disparate sensor greatly improves the detection performance over the single sensor case. Test results of the proposed methods on the two data sets described in Section 2.2 are also provided. The effectiveness of the proposed detection scheme will be presented in terms of probability of detection ( $P_d$ ), probability of false alarm ( $P_{fa}$ ), and the receiver operating characteristic (ROC) curve for various bottom difficulty conditions. Additionally, a comparison of the single sonar and dual disparate sonar cases will be provided in terms of their average false alarms

per image.

The outline of this chapter is as follows. Section 4.2 opens the chapter with the formulation of two channel hypothesis Gauss-Gauss detection test. Section 4.3 presents the single sensor implementation with the corresponding results and analysis. Section 4.4 presents the dual disparate sensor implementation and the corresponding results together with a comparison with those of the single sonar case. Finally, conclusions are made in Section 4.5.

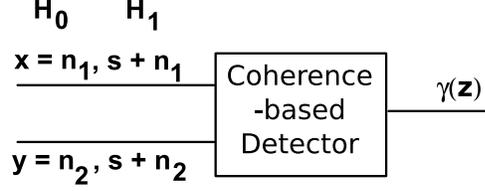
## 4.2 Composite Channel Gauss-Gauss Detection and Formulation [2]

As mentioned before, for the case of detection in sonar imagery using CCA, both channels will contain either background noise or target (signal) and background noise. Figure 4.1 illustrates this hypothesis testing problem and differs from that in Section 3.2.2. The hypotheses are:

$$\begin{aligned} H_0 : \mathbf{z} &= \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix} \\ H_1 : \mathbf{z} &= \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{s} + \mathbf{n}_1 \\ \mathbf{s} + \mathbf{n}_2 \end{bmatrix}. \end{aligned} \quad (4.1)$$

With the change of the hypothesis the new expressions of the log-likelihood and J-divergence need to be formed, in order to correctly apply the Gauss-Gauss detection method to sonar imagery.

For simplicity of the derivations, it is assumed that  $\mathbf{s}$ ,  $\mathbf{n}_1$ , and  $\mathbf{n}_2$  are mutually uncorrelated i.e.  $E[\mathbf{n}_i \mathbf{n}_j^H] = 0$ , and,  $E[\mathbf{s} \mathbf{n}_i^H] = 0$ , and  $E[\mathbf{n}_i \mathbf{n}_i^H] = R_0$  for  $i \neq j$  and  $i, j \in [1, 2]$ . In this case, the covariance matrices  $\bar{R}_0$  and  $\bar{R}_1$  under hypothesis  $H_0$  and  $H_1$ , respectively, yield the following Gauss-Gauss detection test for the composite



**Figure 4.1:** Coherence-based Detector for Composite Channel Hypothesis Case.

observation vector  $\mathbf{z}$ ,

$$\begin{aligned}
 \gamma(\mathbf{z}) &= \begin{cases} 1 \sim H_1, & l(\mathbf{z}) > \bar{\lambda} \\ 0 \sim H_0, & l(\mathbf{z}) \leq \bar{\lambda} \end{cases} \\
 l(\mathbf{z}) &= \mathbf{z}^H \bar{\mathbf{Q}} \mathbf{z} \text{ with } \bar{\mathbf{Q}} = \bar{R}_0^{-1} - \bar{R}_1^{-1}.
 \end{aligned} \tag{4.2}$$

where  $\bar{R}_0$  and  $\bar{R}_1$  are defined as,

$$\bar{R}_0 = \begin{bmatrix} R_0 & 0 \\ 0 & R_0 \end{bmatrix} \tag{4.3}$$

$$\bar{R}_1 = \begin{bmatrix} R_1 = R_s + R_0 & R_s \\ R_s & R_1 = R_s + R_0 \end{bmatrix}. \tag{4.4}$$

Note that we assumed  $E[\mathbf{n}_i \mathbf{n}_i^H] = R_n = R_0$ , for  $i = 1, 2$  and  $E[\mathbf{s} \mathbf{s}^H] = R_s$  where  $R_s$  and  $R_0$  are the covariance matrices under the standard one channel hypothesis that were defined in Section 3.2.1. Also it is assumed that both channels contain uncorrelated noise (background) with covariance matrix  $R_0$ . Clearly this assumption is not true, especially for the dual disparate sensor detector. Nonetheless, it simplifies the derivations significantly.

The expectation of the likelihood  $E[l(\mathbf{z})] = E[\text{tr}(\bar{\mathbf{Q}} \mathbf{z} \mathbf{z}^H)] = \text{tr}(\bar{\mathbf{Q}} E[\mathbf{z} \mathbf{z}^H])$  leads to the J-divergence [2],

$$J = E_{H_1}[l(\mathbf{z})] - E_{H_0}[l(\mathbf{z})] \tag{4.5}$$

$$\begin{aligned}
 &= \text{tr}(\bar{\mathbf{Q}} \bar{R}_1) - \text{tr}(\bar{\mathbf{Q}} \bar{R}_0) \\
 &= \text{tr}(-2\mathbf{I} + \bar{R}_0^{-1} \bar{R}_1 + \bar{R}_1^{-1} \bar{R}_0),
 \end{aligned} \tag{4.6}$$

Using (4.4), we can relate the J-divergence in (4.6) to  $R_0$  and  $R_1$  or to the “signal-to-noise” ratio matrix  $S$  from the standard Gauss-Gauss detector (see Section 3.2.1). Let us first express  $tr(\bar{R}_0^{-1}\bar{R}_1)$  in (4.6) as,

$$tr(\bar{R}_0^{-1}\bar{R}_1) = 2tr(R_0^{-1}R_1) = 2tr(S) = 2tr(\Lambda) \quad (4.7)$$

where  $S$  was defined in (3.18) with eigenvalue decomposition (3.21). Next, we express  $tr(\bar{R}_1^{-1}\bar{R}_0)$  and rewrite it as,

$$tr(\bar{R}_1^{-1}\bar{R}_0) = \begin{bmatrix} R_1 & R_s \\ R_s & R_1 \end{bmatrix}^{-1} \begin{bmatrix} R_0 & 0 \\ 0 & R_0 \end{bmatrix} \quad (4.8)$$

where  $R_1 = R_s + R_0$ . Before we take the inverse of  $\bar{R}_1$ , we decompose it as,

$$\begin{aligned} \begin{bmatrix} R_1 & R_s \\ R_s & R_1 \end{bmatrix} &= \begin{bmatrix} R_1 & R_1 \\ R_1 & R_1 \end{bmatrix} - \begin{bmatrix} 0 & R_0 \\ R_0 & 0 \end{bmatrix} \\ &= - \left( \begin{bmatrix} 0 & R_0 \\ R_0 & 0 \end{bmatrix} - \begin{bmatrix} I \\ I \end{bmatrix} R_1 \begin{bmatrix} I & I \end{bmatrix} \right) \end{aligned} \quad (4.9)$$

Now, using the matrix inversion lemma [51] the inverse of  $\bar{R}_1$  in (4.8) becomes

$$- \left( \begin{bmatrix} 0 & R_0^{-1} \\ R_0^{-1} & 0 \end{bmatrix} - \begin{bmatrix} 0 & R_0^{-1} \\ R_0^{-1} & 0 \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix} \left( -R_1^{-1} + \begin{bmatrix} I & I \end{bmatrix} \begin{bmatrix} 0 & R_0^{-1} \\ R_0^{-1} & 0 \end{bmatrix} \begin{bmatrix} I \\ I \end{bmatrix} \right)^{-1} \begin{bmatrix} I & I \end{bmatrix} \begin{bmatrix} 0 & R_0^{-1} \\ R_0^{-1} & 0 \end{bmatrix} \right) \quad (4.10)$$

Using the fact that  $\begin{bmatrix} 0 & R_0^{-1} \\ R_0^{-1} & 0 \end{bmatrix} \begin{bmatrix} R_0 & 0 \\ 0 & R_0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ , we can now rewrite (4.8) as

$$tr(\bar{R}_1^{-1}\bar{R}_0) = -tr \left( \begin{bmatrix} R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1} & R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1} + I \\ R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1} + I & R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1} \end{bmatrix} \right). \quad (4.11)$$

Again, using the matrix inversion lemma we can write  $(R_1^{-1} - 2R_0^{-1})^{-1} = R_1 - R_1(R_1 - \frac{1}{2}R_0)^{-1}R_1$ . Thus, the diagonal blocks in (4.11) can be rewritten as  $[R_0^{-1}R_1 - R_0^{-1}R_1(I - \frac{1}{2}R_1^{-1}R_0)^{-1}]$ , leading to

$$\text{tr}(\bar{R}_1^{-1}\bar{R}_0) = -2\text{tr}(R_0^{-1}R_1 - R_0^{-1}R_1(I - \frac{1}{2}R_1^{-1}R_0)^{-1}) = -2\text{tr}(S - S(I - \frac{1}{2}S^{-1})^{-1}). \quad (4.12)$$

Now, using (4.7) and (4.12) the J-divergence in (4.6) becomes

$$J = \text{tr}(-2\mathbf{I} + 2S(I - \frac{1}{2}S^{-1})^{-1}). \quad (4.13)$$

Using the fact that S has an eigenvalue decomposition with eigenvalue matrix  $\Lambda$ , (4.13) reduces to

$$J = \text{tr}(-2\mathbf{I} + 2\Lambda(I - \frac{1}{2}\Lambda^{-1})^{-1}) = \sum_{i=1}^N -2 + \frac{4\lambda_i^2}{2\lambda_i - 1}. \quad (4.14)$$

Since the eigenvalues of S,  $(\lambda_i)$ , are related [29] to the squared canonical correlations  $(k_i^2)$  by  $\lambda_i = \frac{1}{1-k_i^2}$  we can rewrite the new J-divergence in terms of the canonical correlations  $k_i^2$ 's, which is a measure of separability of the two hypothesis, as

$$J = \sum_{i=1}^N -2 + \frac{4}{1 - k_i^4}. \quad (4.15)$$

To express the log-likelihood in (4.2) in terms of  $R_0$  and  $R_1$  we can write  $\bar{Q} = (\bar{R}_0^{-1} - \bar{R}_1^{-1})$  and plug in the result for  $\bar{R}_1^{-1}$

$$\bar{Q} = \begin{bmatrix} R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} & R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} \\ R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} & R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} \end{bmatrix}. \quad (4.16)$$

Thus the log-likelihood becomes

$$l(\mathbf{z}) = \mathbf{z}^H \begin{bmatrix} R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} & R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} \\ R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} & R_0^{-1} + R_0^{-1}(R_1^{-1} - 2R_0^{-1})^{-1}R_0^{-1} \end{bmatrix} \mathbf{z}. \quad (4.17)$$

If we choose  $\boldsymbol{\eta} = \begin{bmatrix} R_0^{-1/2} & 0 \\ 0 & R_0^{-1/2} \end{bmatrix} \mathbf{z}$ , then

$$l(\boldsymbol{\eta}) = \boldsymbol{\eta}^H \begin{bmatrix} \mathbf{I} + (\mathbf{S}^{-1} - 2\mathbf{I})^{-1} & \mathbf{I} + (\mathbf{S}^{-1} - 2\mathbf{I})^{-1} \\ \mathbf{I} + (\mathbf{S}^{-1} - 2\mathbf{I})^{-1} & \mathbf{I} + (\mathbf{S}^{-1} - 2\mathbf{I})^{-1} \end{bmatrix} \boldsymbol{\eta}. \quad (4.18)$$

The log-likelihood is now rewritten in terms of  $\mathbf{S}$  and  $R_0$  from the original problem in Section 3.2.2.

Now expressing  $\mathbf{I} + (\mathbf{S}^{-1} - 2\mathbf{I})^{-1}$  in terms of the squared coherence matrix  $\mathbf{C}\mathbf{C}^H = R_0^{H/2} R_1^{-1} R_0^{1/2}$  yields,

$$l(\boldsymbol{\eta}) = \boldsymbol{\eta}^H \begin{bmatrix} 1/2\mathbf{I} - 1/4((\mathbf{C}\mathbf{C}^H)^{-1} - 1/2\mathbf{I})^{-1} & 1/2\mathbf{I} - 1/4((\mathbf{C}\mathbf{C}^H)^{-1} - 1/2\mathbf{I})^{-1} \\ 1/2\mathbf{I} - 1/4((\mathbf{C}\mathbf{C}^H)^{-1} - 1/2\mathbf{I})^{-1} & 1/2\mathbf{I} - 1/4((\mathbf{C}\mathbf{C}^H)^{-1} - 1/2\mathbf{I})^{-1} \end{bmatrix} \boldsymbol{\eta}. \quad (4.19)$$

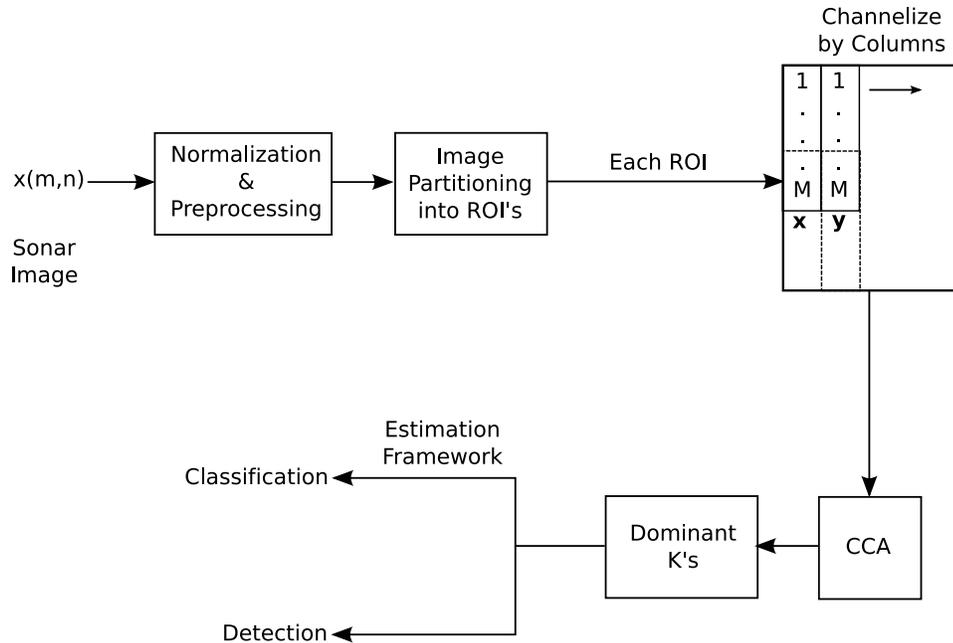
which can further be expressed in term of canonical correlation matrix  $\mathbf{K}$  as

$$l(\boldsymbol{\eta}) = \boldsymbol{\eta}^H \begin{bmatrix} \mathbf{F} & 0 \\ 0 & \mathbf{F} \end{bmatrix} \begin{bmatrix} 1/2\mathbf{I} - 1/4(\mathbf{K}^{-2} - 1/2\mathbf{I})^{-1} & 1/2\mathbf{I} - 1/4(\mathbf{K}^{-2} - 1/2\mathbf{I})^{-1} \\ 1/2\mathbf{I} - 1/4(\mathbf{K}^{-2} - 1/2\mathbf{I})^{-1} & 1/2\mathbf{I} - 1/4(\mathbf{K}^{-2} - 1/2\mathbf{I})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{F}^H & 0 \\ 0 & \mathbf{F}^H \end{bmatrix} \boldsymbol{\eta}. \quad (4.20)$$

Using the SVD of the coherence matrix  $\mathbf{C} = R_0^{H/2} R_1^{-H/2} = \mathbf{F}\mathbf{K}\mathbf{G}^H$ , it is easy to show [29] that  $\mathbf{F}^H R_0^{-1/2} = \mathbf{K}^{-1} \mathbf{G}^H R_1^{-1/2}$  and the log-likelihood can be rewritten in terms of the canonical coordinates as

$$l(\mathbf{w}) = \mathbf{w}^H \begin{bmatrix} 1/2\mathbf{I} - 1/4(\mathbf{K}^{-2} - 1/2\mathbf{I})^{-1} & 1/2\mathbf{K}^{-1} - 1/4(\mathbf{K}^{-1} - 1/2\mathbf{K})^{-1} \\ 1/2\mathbf{K}^{-1} - 1/4(\mathbf{K}^{-1} - 1/2\mathbf{K})^{-1} & 1/2\mathbf{K}^{-2} - 1/4(\mathbf{I} - 1/2\mathbf{K}^2)^{-1} \end{bmatrix} \mathbf{w}, \quad (4.21)$$

where  $\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}$  is the composite canonical coordinate vector [27] which is related to the canonical coordinate vector  $\mathbf{v}$  in Section 3.3. Hence, the log-likelihood ratio is only expressed in terms of the canonical coordinates  $\mathbf{u}$  and  $\mathbf{v}$  and the corresponding correlations  $k_i$ 's.



**Figure 4.2:** Block Diagram of the Single-Sensor CCA-based Detection Method.

### 4.3 Single-Sonar Implementation and Results

The first application of the proposed coherence-based composite channel detector is the single sonar case which provides one image of the ocean floor. In this case, CCA provides an ideal framework for simultaneous coherent detection and feature extraction of target attributes that are present in an ROI within the image. In this situation both channels of the CCA are formed from one ROI within the sonar image and for every ROI we compute canonical coordinates and correlations to form the log-likelihood function (4.21) in order to determine whether the ROI contains a target or background only. Our *detection hypothesis* in this single sonar coherence analysis is that presence of objects in an ROI leads to high level of coherence measure comparing to that of the background clutter only. The block diagram of the proposed single-sonar target detection and feature extraction system is presented in Figure 4.2.

### 4.3.1 System Overview

In order to prepare the Sonar8 data for CCA, the images are first normalized using a serpentine forward-backward filter that was described in Section 2.3.1. The purpose of this normalization is to help distinguish the target's highlight and shadow signature from the bottom and artifacts present in the image.

After the normalization process, the first 325 pixels are ignored which depends on the sonar altitude as it traveled through the water column. This corresponds to  $1/10^{th}$  of the maximum range of the sonar. Next, the image is partitioned into overlapping ROI's of size  $M \times N$ . For the Sonar8 data set the ROI size of  $46 \times 180$  was experimentally determined to be optimal considering the average size and shape of the targets in this data set. The overlap along the horizontal and vertical directions was 50% in order to ensure that a target would be covered by more than one ROI in case of splitting. Each ROI is then channelized in a column-wise fashion for the two-channel CCA. The  $\mathbf{x}$  and  $\mathbf{y}$  channels consist of the first 30 pixels in one column  $\mathbf{x}$  and the first 30 pixels in the adjacent column  $\mathbf{y}$ . This process is continued moving in the horizontal direction across the ROI. On the next pass through the ROI, the channels are given a 50% overlap in the vertical direction to ensure complete coverage of the target in the ROI (see Figure 4.2). The idea behind this channelization is to look for common coherent attributes that can be used to relate one channel to the other and to capture the target structure. Clearly, for background ROI's high level of coherence among consecutive columns (channels) does not exist. In this development we use the dominant canonical correlations which hold the coherent information between the two channels. One may also use the subdominant canonical correlations in the detection framework to detect incoherence (or change) between the two channels.

After the channelization, the canonical coordinates and correlations are found for each interrogated ROI. From the canonical correlations the J-divergence measure is calculated and compared against a threshold. Any ROI whose J-divergence falls

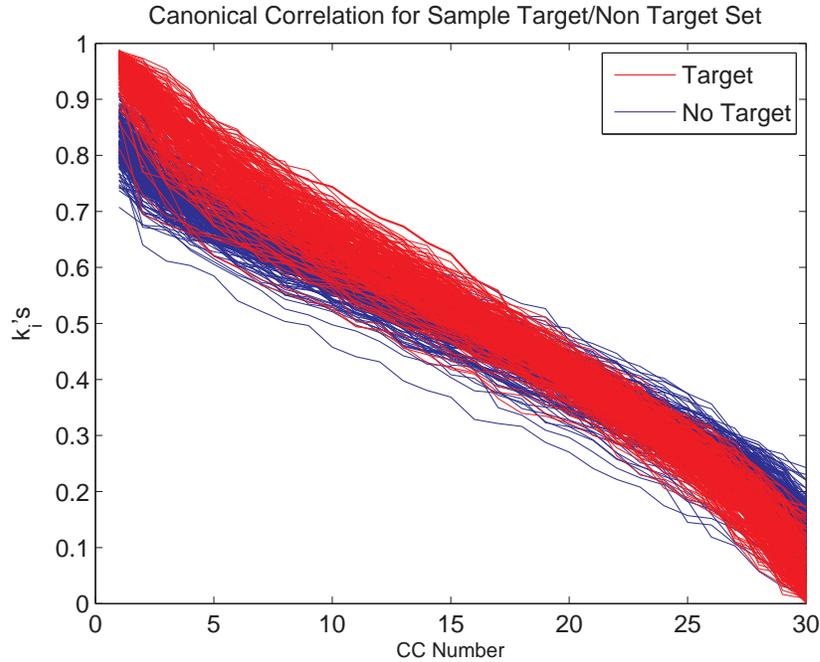
above the threshold is flagged for further investigation. Next, the log-likelihood ratio expression is found for each column pair within the ROI. A detection score is then created based on the percent of the log-likelihood measurements within the ROI that fall above the detection threshold. Detection score  $\geq 50\%$  signifies the presence of a target within the current ROI.

### 4.3.2 Results and Observations

The single-sonar detector was applied to the Sonar8 data set. A detailed description of the data set was provided in Section 2.2.1. The data set contains 137 images containing 286 targets with some of the images containing more than one target. The data set is broken up into easy, medium, and hard cases depending on the density of bottom clutter.

To show the separability of the dominant canonical correlations for ROI's that contain targets plus background and those that contain only background, a test was conducted on the entire set of target ROI's and a random set of ROI's containing mainly background (for all three bottom cases) of the same number of target ROI's. The plots of the 30 canonical correlations of ROI's containing targets and those containing background only are shown in Figure 4.3. As can be seen, there is good separation, especially for dominant canonical correlations, between targets and background, which can be attributed to the greater coherence in the channels across the targets versus the background.

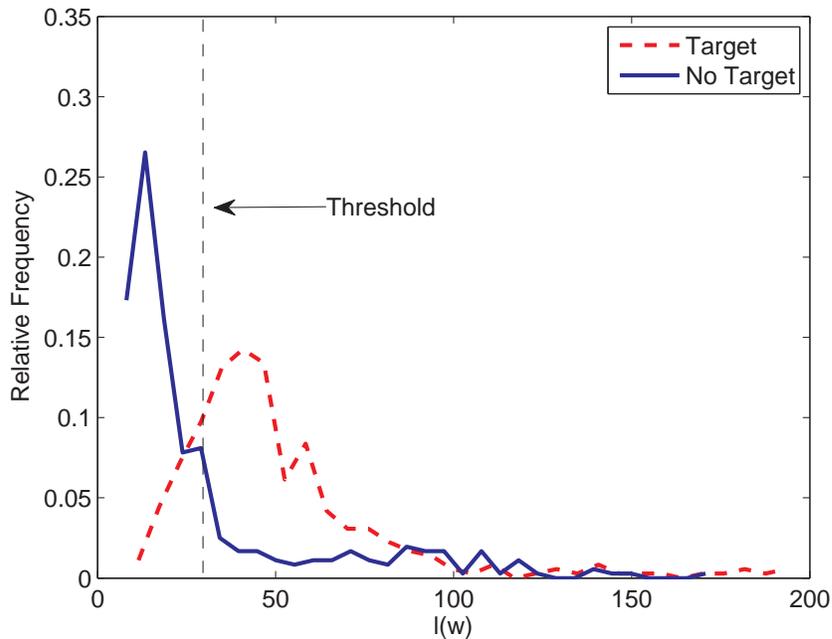
Figure 4.4 shows the histogram of the log-likelihood ratio values of one target ROI and one random background ROI from the data set. From the example ROI we see that we have good separation between the two hypotheses which in turn allows for good detection. Using the entire set of target and the same number of random background ROI's, an optimum threshold was determined to be 38.2. This detection threshold is represented by the vertical dashed line in Figure 4.4. This detection



**Figure 4.3:** Plot of Canonical Correlations for Target and Background for Sonar8.

threshold is the same across the easy, medium, and hard data sets as all three sets exhibited similar histogram plots with the similar separation amounts.

The system is then implemented on the entire Sonar8 data set provided by the NSW-C-Panama City. The coherence-based detector performs well on the easy and medium set, but falls short on the hard set. More specifically, for the easy cases, there are 106 images containing 233 targets and the detector detected 229 targets with an average of 19 false alarms per image. For the medium cases, there are 29 images containing 49 targets and the detector detected 43 targets with an average of 22 false alarms per image. Lastly, for the hard cases, there are 2 images, containing 4 targets and the detector detected 3 targets with an average of 30 false alarms per image. As expected, higher false alarms are generated as the clutter density increased, due to the highly textured and dense nature of the bottom clutter. Out of the 286 targets in the entire data set only 11 targets were missed. The miss-detected targets were those that had little target structure or were obscured by very dense clutter. When

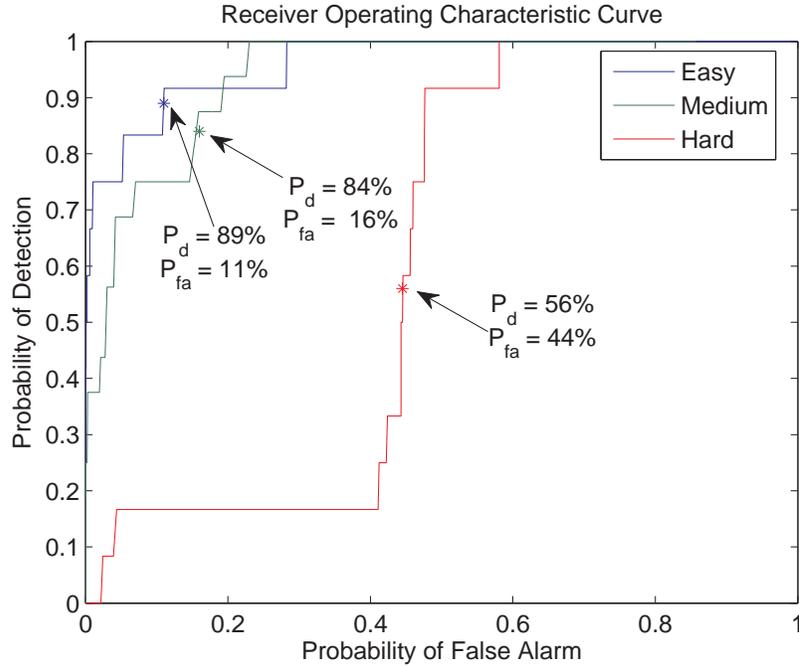


**Figure 4.4:** Histogram of Example Log-Likelihood Values for Target and Background for Sonar8.

the target structure is reduced there is little or no shadow structure leaving mainly a bright spot in the ROI hence making the detection difficult.

The ROC curves for the three different subsets of the Sonar8 database are presented in Figure 4.5. At the knee point for the easy case we have  $P_d = 89\%$  and  $P_{fa} = 11\%$ . For the medium case the knee point exhibits  $P_d = 84\%$  and  $P_{fa} = 16\%$  and for hard case the knee point gives  $P_d = 56\%$  and  $P_{fa} = 44\%$ , which is poor due to the fact only a few images existed in this subset. Clearly, it is hard to determine the detector’s performance when only a few target cases are available. Overall, the detector performs well on the easy and medium cases with a good probability of detection and a low false alarm rate. For the hard cases the detector suffers, due to the large amount of background noise and bottom clutter.

These results motivate the use of dual or multiple disparate sonar platforms. By using the coherent information across two or more disparate sonar imagery we can greatly reduce the false alarm rate and increase the detection performance by

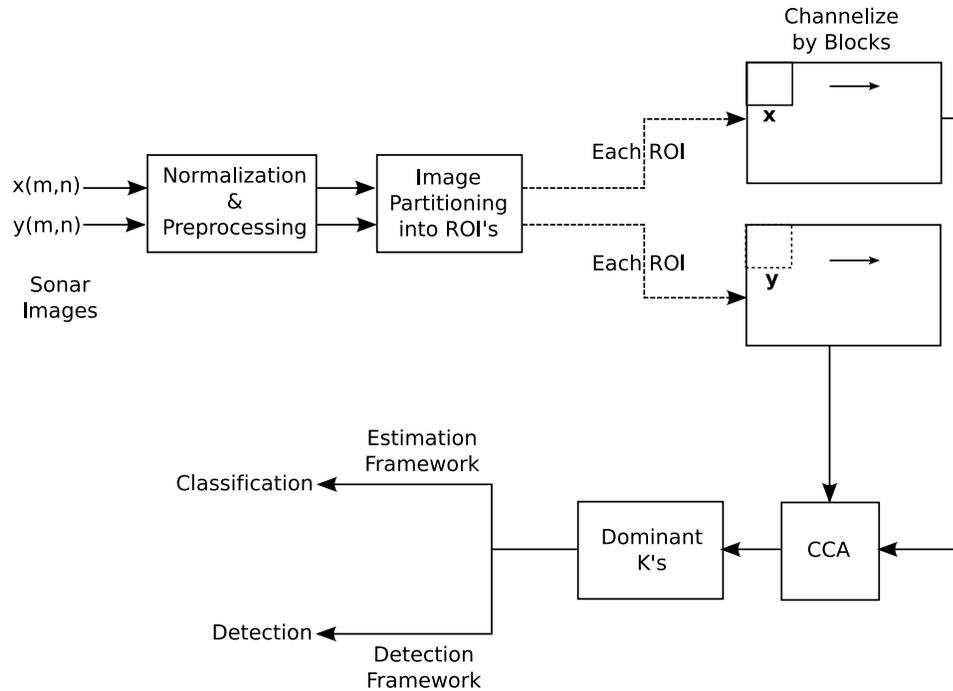


**Figure 4.5:** ROC Curves for the Sonar8 Data Set.

capturing the target information at different sonar frequencies, beamwidth, resolution, and perhaps location, even in difficult bottom conditions.

## 4.4 Dual Disparate Sensor Implementation and Results

In this section, the coherence-based detector, is implemented on dual disparate sonars where the disparateness is in the operating frequency, bandwidth, and beamwidth (resolution). This case is of particular interest as it allows coherent information between two different sensors looking at the same region to be used to make a detection decision. There are several key benefits that dual disparate underwater target system offers comparing to the single sensor system. These include: (a) the ability to collaboratively detect and agree on the events observed by different sensors, (b) the ability to generate a set of coherent features extracted among all the data of the sensors instead of extracting feature sets from each individual sensory data, and (c)



**Figure 4.6:** Coherent-Band detection and feature extraction from Dual Disparate Sonar Platform.

the ability to reduce false alarm caused by man-made and natural clutter by using different types of data observed by different sensors.

For dual sonar platforms, the CCA data channels consist of a pair of blocks within ROI's that are co-registered over the same region. As in the previous case, CCA provides an ideal framework for simultaneous coherent detection and feature extraction of target attributes present in both images. A block diagram of the dual disparate sensor implementation is provided in Figure 4.6. These extracted canonical correlations for each pair of ROI's provide a coherence (or incoherence) measure that can be used to determine if a target is present (or absent) in the processed ROI's. Our *detection hypothesis* in this dual-sensor coherence analysis is that presence of objects in a pair of co-registered ROI's leads to high level of coherence measure comparing to that of the background clutter only.

#### 4.4.1 System Overview

The CCA-based detector is applied to a dual sonar data set consisting of high-resolution side-looking sonar images as well as broadband sonar images, the data set description was given in Section 2.2.2 and is referred to as the multi-platform sonar data set. The pair of sonar images, which cover the same area with possibly different spatial resolution and contrast, form the two-channels in the CCA processing. Although different sonar platforms observe the same area at different elevation, grazing angle, beamwidth, and frequency characteristics, the target returns are more coherent than those of the background (*detection hypothesis*). CCA can isolate and represent these coherent (or common) features between the two data channels. Once the common features between the two sonar images are identified, the *coherence pattern* represented by the dominant canonical correlations can be used for subsequent classification (*classification hypothesis*) as the coherent information is captured in only few dominant canonical correlations corresponding to the dominant canonical coordinates [29].

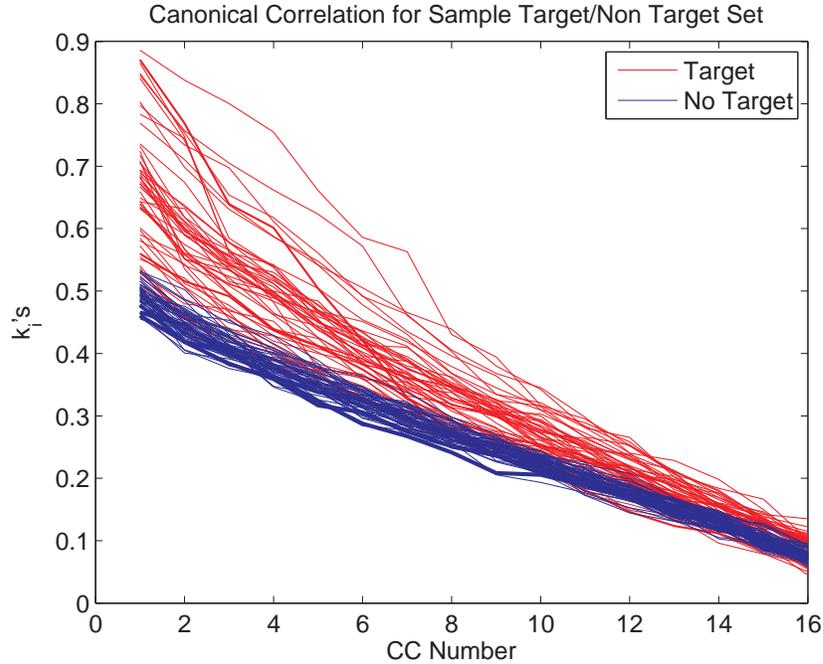
The high-frequency (HF) side-scan sonar produces high resolution imagery in which targets typically have good definitions with prominent highlights and shadows, though clutter is also prominent. In the broadband (BB) sonar images the target highlight still shows up while clutter is not as prominent. Thus, the idea behind using two-channel CCA is to exploit the high coherence between the two sonar images when targets are present. As mentioned before, using CCA the canonical correlation features that capture this coherence property remain somewhat unchanged to perspective transformation on the images, which could be caused by variations in grazing angle and object orientation with respect to the sonar platforms. Moreover, the beamformed complex-valued data of the sonar systems are used. Clearly, the capability to process complex-valued data and retain valuable phase information is another benefit of the CCA framework.

The multi-platform sonar data set contains 59 images containing 53 targets with some of the images containing more than one target. For the present study, the optimum ROI sizes for the HF and BB sonar imagery were experimentally chosen to be  $72 \times 112$  pixels and  $24 \times 224$ , respectively. The images are then partitioned into ROI's with a 50% overlap along both range and cross-range directions. Each ROI is then channelized by a rectangular blocking scheme with the dimension of the blocks for the HF and BB sonar images being  $6 \times 4$  and  $2 \times 8$  pixels, respectively. These blocks from realization of the two channel data in the CCA process. Then, canonical correlations are computed for each pair of ROI's using these samples. These canonical correlations can be used for both detection and classification as mentioned previously.

Next, for the detection process, the log-likelihood ratio expression in (4.21) is computed for each block within the ROI pair. A detection score is then created based on the percent of the log-likelihood measurements within ROI pair that fall above the detection threshold. Detection score  $\geq 50\%$  signifies presence of a target within that pair of ROI's.

#### 4.4.2 Results and Observations

To show the separability of the dominant canonical correlations for ROI's that contain targets over background and those that contain only background, a test was conducted on the entire set of 53 target ROI's and a same size randomly selected set of ROI's containing only background clutter. The plots of 16 canonical correlations of ROI's containing targets and those containing background only are shown in Figure 4.7. As can be seen, there is good separation between targets and background, especially for dominant canonical correlations. This can be attributed to the greater coherence between  $\mathbf{x}$  and  $\mathbf{y}$  channels across the target ROI's versus those over background clutter where there is more randomness. In other words, there is greater coherence between the HF and BB sonar images when a target is present in both ROI's versus the case

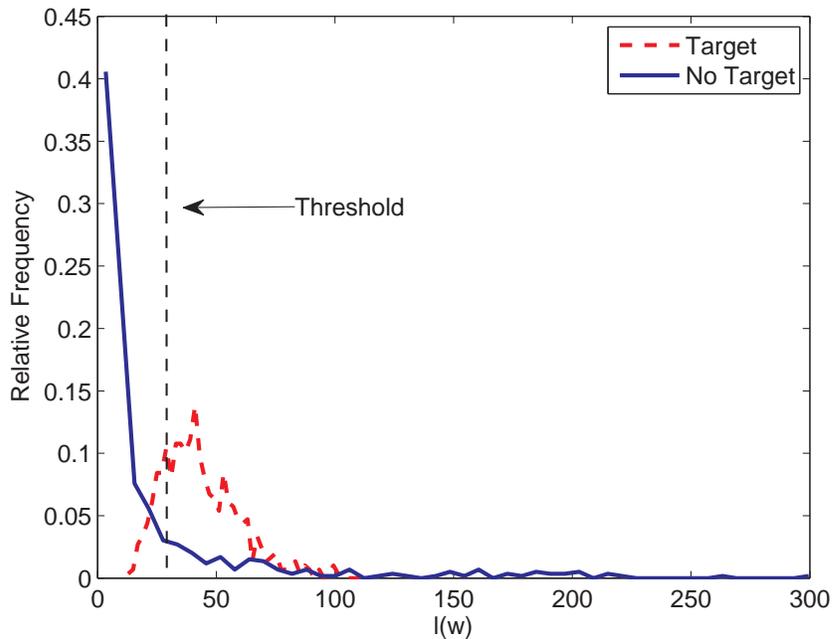


**Figure 4.7:** Plot of Canonical Correlations for Target and Background for HF-BB.

where both ROI's contain only background noise.

Figure 4.8 shows the histogram of the log-likelihood ratio values of one target ROI and one random background ROI in the HF-BB detector. Using the entire set of target and background ROI's, an optimum threshold was determined to be 26.4 for the detector. The threshold is shown by dotted vertical lines in Figure 4.8. These values were then used as detection thresholds for all the ROI's in all the images in the data set. That is, any block within an ROI pair whose log-likelihood ratio falls above the relevant threshold is flagged as a target block for that ROI pair.

The detector was executed on the entire data set with good results. The detector detected 51 of the 53 targets with an average of 10.13 false detections per image. The ROC curve for the HF-BB detector is presented in Figure 4.9. At the knee point of the ROC curve for the HF-BB detector we have  $P_d = 95\%$  and  $P_{fa} = 5\%$ . It was observed that for those targets that were missed the coherence was low due to the fact that the target only appeared in one of the images. Overall, the detector performed

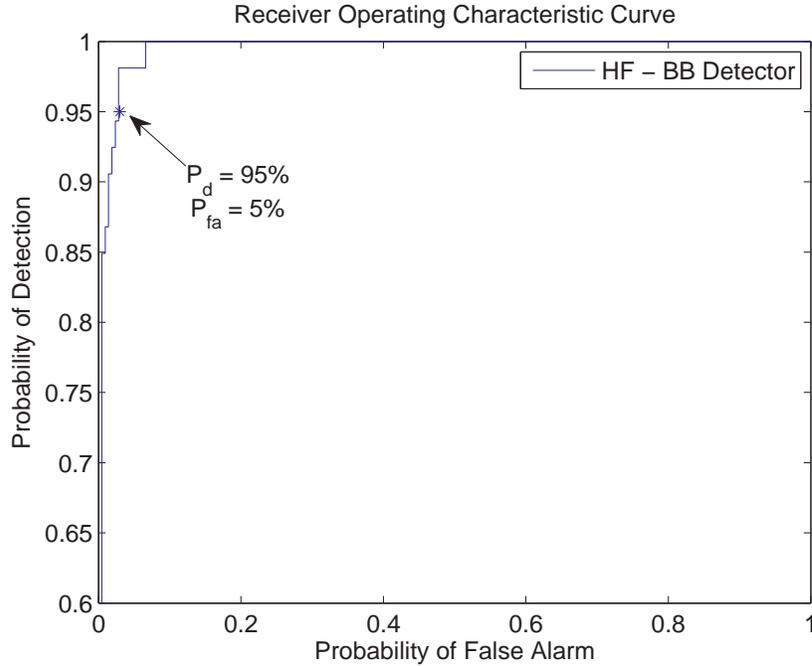


**Figure 4.8:** Histogram of Example Log-Likelihood Values for Target and Background for HF-BB.

extremely well given the small number of targets and non-targets used to form the detection threshold.

## 4.5 Conclusion

In this chapter, a new composite two-channel coherence-based detector was developed for single sensor and dual disparate sensory systems. The single sensor utilizes one high frequency side-scan sonar image while the dual disparate sensor utilizes two side-scan sonar images captured at disparate frequency and resolution. Due to the nature of target detection in both cases the resulting hypotheses for the detection problem contain two channels that both capture target or background. This significantly deviates from the standard detection problems where both hypotheses contain one channel only. Thus, new formulations for the log-likelihood ratio and J-divergence are needed to account for these major differences. These formulations were applied



**Figure 4.9:** ROC Curve for the HF-BB Detector.

to the single sonar case where the goal is to determine if a target is present in an ROI by checking to see if consecutive columns within a ROI contain either signal plus noise or noise only. The formulation was then applied to the dual disparate sonar case where the goal is to determine if a target is present in a pair of co-registered ROI's by checking to see if pairs of blocks within ROI's contain either signal plus noise or noise only.

The single sonar results on the Sonar8 data set demonstrated that 275 out of the 286 targets were successfully detected with an average of 24 false alarms per image. As far as performance in different clutter densities is concerned, the ROC curve for the easy cases (low clutter density) showed  $P_d = 89\%$  and  $P_{fa} = 11\%$  at the knee point. For the medium clutter density the knee point exhibited  $P_d = 84\%$  and  $P_{fa} = 16\%$  and finally for the hard cases (very difficult and high density clutter) the knee point showed  $P_d = 56\%$  and  $P_{fa} = 44\%$ .

The use of multiple platform sonar allows one to use a high resolution high frequency sonar with good target definition and low resolution broadband sonar with good clutter suppression abilities to improve the detection performance over that of the single sonar platform. For the dual disparate case the detector detected 51 out of the 53 targets and had an average of 10 false alarms per image. The knee point on the ROC showed  $P_d = 95\%$  and  $P_{fa} = 5\%$  for the HF-BB detector. This indicates that the combination of the information from the BB and HF sonars resulted in much better detection performance when compared to the use of a HF sonar only. The development of the two channel coherence-based detector and its ease in implementation in both situations together with its great detection performance make this approach very valuable in underwater target detection problems.

In Chapter 5 new distributed collaborative detection system is developed where two dual disparate detectors are used as local decision makers. The final decision is generated by the fusion center and is based on the individual decisions of the local decision makers and an observation made by the fusion center. The corresponding fusion rule will be derived and the results will be compared to those provided in this chapter.

# CHAPTER 5

## MULTI-PLATFORM DISTRIBUTED DETECTION IN SONAR IMAGERY

### 5.1 Introduction

To optimize the overall detection performance and increase the practical feasibility of the coherence analysis framework in realistic multiple disparate sonar platform systems the use of multiple decision makers is a necessity. In this chapter, we propose a novel distributed detection system using the coherence-based detector developed in the previous chapter. In this system, there is a group of  $N$  local decision-makers used to produce  $N$  separate decisions. Each local decision maker performs coherence-based detection and feature extraction based upon two sonar images and generates local decisions and confidence measures as well as a set of canonical correlation features extracted within the ROI's of that pair of sonar images. Each decision maker then sends its decision to a fusion center where a final decision is made based upon the preliminary decisions and the fusion center's own observation of the environment. The final decision is made based on a likelihood ratio test [48] that utilizes not only the conditional probabilities of the local decisions but also the conditional probabilities of the fusion center's own observation.

The developed distributed detection system is implemented using two (or more) dual-platform detection systems that were described in Section 4.4 to generate two local decisions that are subsequently processed by the fusion center. The motivation behind this setup is that by using multiple autonomous underwater vehicles (AUV's) each equipped with two sensors and a local detection and decision-maker one can

perform target detection and classification using multiple ( $> 2$ ) platforms. Moreover, for the final decision, the amount of data that has to be sent to the fusion center is drastically reduced as only the local decisions together with simple statistical attributes extracted from the ROI's need to be sent hence making this system amenable to real-time network-centric sensor analysis (NSA) scenarios. The fusion of the individual decisions leads to higher probability of detection and the lower false alarm rate due to the fact that more multiple decisions are made about the observations from the environment.

Test results of the proposed method on the multi-platform sonar data set described in Section 2.2 are provided. The effectiveness of the proposed distributed detection system will be presented in terms of probability of detection ( $P_d$ ), and probability of false alarm ( $P_{fa}$ ). Additionally, a comparison of the performance of the individual local decision makers will be made, especially in terms of the number of false alarms per image.

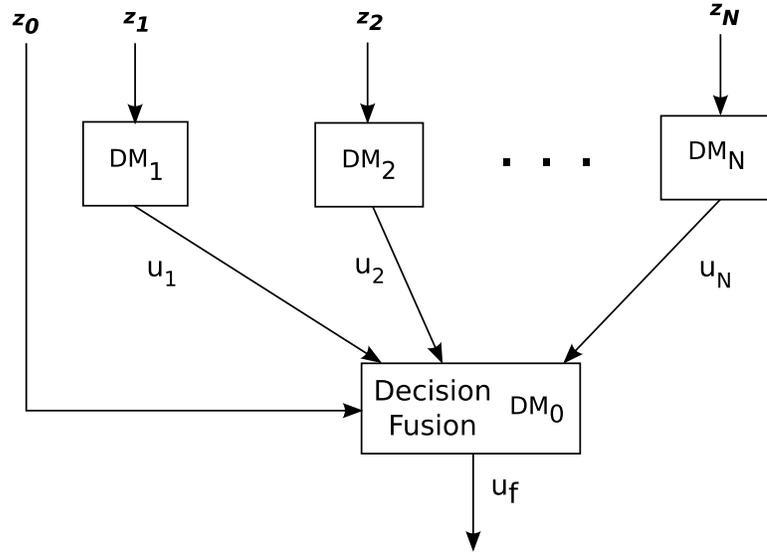
The outline of this chapter is as follows. Section 5.2 introduces the distributed detection system and the development of the corresponding fusion rule. Section 5.3 presents the implementation of the distributed detection system and corresponding results. A discussion of the results and a comparison to the single dual disparate case in Chapter 4.4 is also provided. Finally, conclusions are made in Section 5.4.

## 5.2 Distributed Detection

Lets now consider the distributed binary hypothesis testing problem where a decision must be made between  $H_0$  and  $H_1$  at  $N$  local decision makers. Figure 5.1 illustrates the structure of such a system consisting of  $N$  local decision makers denoted  $DM_i$  each making its own independent observation  $\mathbf{z}_i$  from the environment. The prior probabilities of the two hypotheses  $H_0$  and  $H_1$  are  $P_0$  and  $P_1$ , respectively. Each local decision maker then receives a composite observation vector  $\mathbf{z}_i$ ,  $i \in [1, N]$  and

the set of decisions  $\mathbf{z}_1, \dots, \mathbf{z}_N$  are assumed to be conditionally independent with known conditional densities  $p(\mathbf{z}_i|H_j)$ ,  $i \in [1, N]$ ,  $j = 0, 1$ . The local decision makers then send their decisions  $u_i$  to the fusion center where based upon the received local decisions and on its own observation of the environment  $\mathbf{z}_0$ , the fusion center makes a final decision  $u_f$ .

Let us denote the decision rule of the fusion center as  $\gamma_f(\cdot)$ , where  $u_f = \gamma_f(u_1, \dots, u_N, \mathbf{z}_0) \in [0, 1]$  and denote  $\gamma_i(\cdot)$  as the decision rule of the local decision maker where  $u_i = \gamma_i(\mathbf{z}_i)$ . Furthermore, let us assume that the overall cost is a function of only the final decision,  $u_f$  and the true hypothesis. In other words, let  $C_{u_f, H_j}$  be the cost of decision  $u_f$  by the fusion center when hypothesis  $H_j$  is true. We will make the assumption that the cost of making erroneous decisions is greater than the cost of a correct decision,  $C_{01} > C_{00}$  and  $C_{10} > C_{11}$ . Therefore we need to find the decision rule  $\gamma_f$  such that the expected cost  $C = E[C_{u_f, H_j}]$  is minimized.



**Figure 5.1:** Block Diagram of the Distributed Detection Problem.

If we assume that the decision rules of the local decision makers are given, then the optimal strategy for the fusion rule is a likelihood ratio test as the problem is

simply a centralized detection problem and is a result of the centralized Neyman-Pearson Lemma [50]. Here we assume that the decisions  $u_i$ 's and  $\mathbf{z}_0$  are conditionally independent, and therefore the likelihood ratio test is given by [48],

$$\gamma_f(u_1, \dots, u_N, \mathbf{z}_0) = \frac{p(\mathbf{z}_0|H_1) \prod_{i=1}^N p(u_i|H_1)}{p(\mathbf{z}_0|H_0) \prod_{i=1}^N p(u_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \quad (5.1)$$

where  $\lambda$  is the decision threshold (3.8) and is defined as

$$\lambda = \frac{P_0[C_{10} - C_{00}]}{P_1[C_{01} - C_{11}]} \quad (5.2)$$

For the local decision makers, since we assumed that the observations are conditionally independent and if we assume the fusion rule  $\gamma_f(\cdot)$  is known then the optimal decision rule,  $\gamma_i(\cdot)$  for the  $N$  local decision makers is a likelihood ratio test given by [47], [48]

$$\gamma_i(\mathbf{z}_i) = \frac{p(\mathbf{z}_i|H_1)}{p(\mathbf{z}_i|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_i \quad (5.3)$$

where  $\eta_i$  is the local decision threshold. Therefore, by using our coherence-based detector as the local decision maker, the fusion rule above will be optimal as the coherence-based detector is a likelihood ratio test.

In order to find the conditional densities,  $p(\mathbf{z}_0|H_j)$  and  $p(u_i|H_j)$  for the fusion rule, one may use a density estimator such as the probabilistic neural network (PNN) or back-propagation neural network (BPNN) [52] - [54]. To make the fusion rule more suitable for the approach of using neural network density estimators one may use Bayes' rule to rewrite the fusion rule in terms of the *a posteriori* conditional densities  $p(H_j|\mathbf{z}_0)$  and  $p(H_j|u_i)$  that are generated at the outputs of the PNN/BPNN. This gives,

$$\gamma_f(u_1, \dots, u_N, \mathbf{z}_0) = \frac{p(H_1|\mathbf{z}_0) \prod_{i=1}^N p(H_1|u_i)}{p(H_0|\mathbf{z}_0) \prod_{i=1}^N p(H_0|u_i)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_1^N [C_{10} - C_{00}]}{P_0^N [C_{01} - C_{11}]} \quad (5.4)$$

The PNN, which implements the Parzen non-parametric probability density function estimation [55] and Bayes decision rule can be used to estimate the conditional

densities  $p(H_j|\mathbf{z}_0)$ . The PNN consists of three feedforward layers, the input layer, pattern layer, and summation layer [52]. Feature vectors are applied to the input layer, which passes them to each neuron in the pattern layer. The pattern layer consists of  $K$  pools of pattern neurons, where  $K$  is the number of classes or hypotheses (in this case  $K = 2$ ). In each pool  $k \in [0, K - 1]$ , there are  $N_k$  pattern neurons, each of which represents exactly one pattern from the training set for hypothesis  $H_k$ ,  $k = 0, 1$ . For the input observation vector  $\mathbf{z}_0$  with dimension  $d$ , the output of each pattern layer neuron is

$$f(\mathbf{z}_0; \mathbf{w}_k^{(j)}, \sigma) = \frac{1}{N_k(2\pi)^{d/2}\sigma^d} \exp \left[ -\frac{(\mathbf{z}_0 - \mathbf{w}_k^{(j)})^T(\mathbf{z}_0 - \mathbf{w}_k^{(j)})}{2\sigma^2} \right], \quad (5.5)$$

where  $\mathbf{w}_k^{(j)}$  is the weight vector of the  $j$ th neuron in the  $k$ th pool, and the nonlinear function  $f(\cdot)$  represents the activation functions of the neurons. In the summation layer, the  $k$ th neuron associated with hypothesis  $H_k$ ,  $k = 0, 1$ , forms the weighted sum of all the outputs from the neurons in the  $k$ th pool in the pattern layer. The weights in the summation layer are determined by the decision cost function and the prior hypotheses probabilities. For the “0-1” cost function and equally likely hypotheses, the weights will be one for all the neurons in the summation layer. For the input pattern  $\mathbf{z}_0$  of an unknown class, a final decision is made through a simple comparison of the PNN outputs  $O_k(\mathbf{z}_0)$ , i.e.,

$$\mathbf{z}_0 \in H_1, \quad \text{if } O_1(\mathbf{z}_0) > O_0(\mathbf{z}_0). \quad (5.6)$$

Under certain conditions [52], the outputs of the PNN correspond to the *a posteriori* conditional probabilities, i.e.  $O_1(\mathbf{z}_0) \approx p(H_1|\mathbf{z}_0)$ , when  $\mathbf{w}_1^{(j)} = \mathbf{z}_{0,1}^{(j)}$ , i.e., the weight vector of the  $j$ th neuron in pool 1 is set to the training sample  $\mathbf{z}_{0,1}^{(j)}$  belonging to hypothesis  $H_1$ . Although this training process is very fast, a very large network may potentially be formed, especially if the number of samples in the training set is large.

In order to generate the likelihood ratio  $\frac{p(H_1|\mathbf{z}_0)}{p(H_0|\mathbf{z}_0)}$  from the two PNN outputs for the binary hypothesis problem, we simply use  $\frac{p(H_1|\mathbf{z}_0)}{p(H_0|\mathbf{z}_0)} = \frac{\hat{O}_1(\mathbf{z}_0)}{\hat{O}_0(\mathbf{z}_0)}$ , where  $\hat{\cdot}$  means normalized such that  $\hat{O}_0(\mathbf{z}_i) + \hat{O}_1(\mathbf{z}_i) = 1$ . This accounts for inaccuracies in estimating the conditional densities using the PNN. Thus, the case  $\hat{O}_1(\mathbf{z}_0) = 1$  implies that the fusion center strongly believes  $\mathbf{z}_0 \in H_1$ , while  $\hat{O}_1(\mathbf{z}_0) = 0$  means that the fusion center strongly believes  $\mathbf{z}_0 \in H_0$ . In the case when  $\hat{O}_1(\mathbf{z}_0) = 0.5$ , the fusion center is unsure which hypothesis is true.

The conditional densities  $p(H_j|u_i)$ ,  $i \in [1, N]$ ,  $j = 0, 1$  that represent the confidence in local decision maker's decisions  $u_i$  can be generated using a BPNN. In other words, once the BPNN-based probability estimator has received  $u_i$ , it estimates  $p(H_j|u_i)$ 's such that

$$y_i(j) \approx p(H_j|u_i), \quad i \in [1, N], \quad (5.7)$$

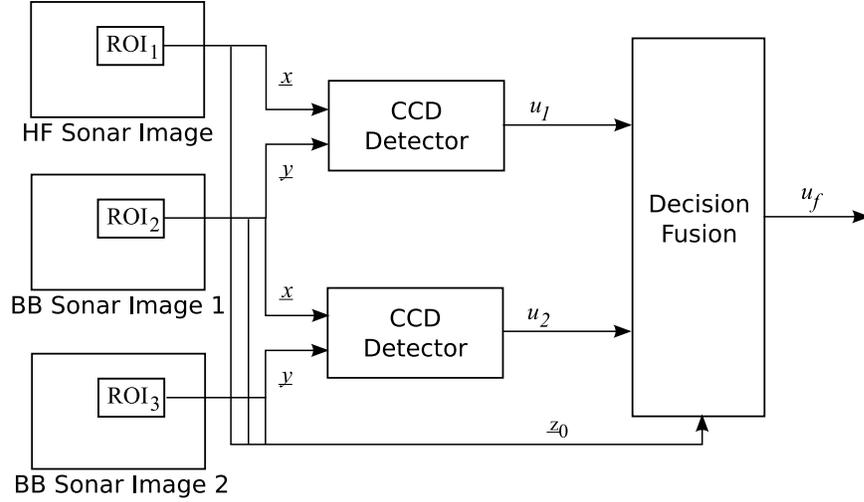
where  $y_i(j)$  is the  $j$ th ( $j \in \{0, 1\}$ ) output of the BPNN for the decision  $u_i$ . The BPNN is trained to capture this mapping between the preliminary decisions in the training data set and the hypothesis. It is well-known [52] that, if properly trained, the BPNN performance approximates an optimal Bayesian estimator.

Using the trained PNN and BPNN the final decision rule can then be implemented and used to make a final decision based on all the local decision makers decisions and the fusion center's own observation,  $\mathbf{z}_0$ , from the environment.

### 5.3 Multi-Platform Distributed Implementation and Results

We now implement the distributed detection system where more than 2 sonar images are used from the multi-platform sonar data set. A system diagram is presented in Figure 5.2. As can be seen in this figure, the first detector uses one HF image and one BB image with a different spatial resolution, while the second detector uses that same BB image and a second BB image with different frequency characteristics.

Both detectors are implemented using the dual platform-based detection presented in Section 4.4 which become the two local decision makers for the distributed set up. Again, our *detection hypothesis* in this distributed dual-sensor coherence analysis is that presence of objects in a pair of co-registered ROI's leads to high level of coherence measure comparing to that of the background clutter only.



**Figure 5.2:** Block Diagram of the Multi-platform Detection System.

### 5.3.1 System Overview

For the each of the local detectors the same data preprocessing and preparation as described in Section 4.4.1 are used here. We will represent the first local decision maker by BB<sub>1</sub>-BB<sub>2</sub> and the second local decision maker by HF-BB<sub>1</sub>. The outputs of each detector is fused to generate a final decision using the fusion rule developed in the previous section. It must be noted that in the first detector, coherence information in the two BB sonar images is exploited to further suppress clutter ROI's and detect potential targets. The coherence information in the second detector solidifies the decision in the first detector by verifying the joint presence in the HF and BB images. The fusion of the decisions using developed fusion rule in (5.4) leads to reduced number of false alarms and misdetections.

The detection process at each of the individual detectors (decision-makers) is the

same as that described in Section 4.4.1. That is, the log-likelihood ratio expression in (4.21) is computed for each block within the ROI pair. A detection score was then created based on the percent of the log-likelihood measurements within the ROI pair that fall above the detection threshold. Detection score  $\geq 50\%$  signifies presence of a target within that pair of ROI's. The detection scores, namely  $u_i$ 's, of the two detectors are then sent as the decisions of the two detectors to the fusion center. At the fusion center the fusion rule takes it's own observation,  $\mathbf{z}_0$  from the three ROI's and the two decisions from the local decision makers to make the final decision,  $u_f$ .

In order to estimate the conditional densities needed for the fusion rule, the two neural networks, namely the PNN and BPNN, need to be trained, validated, and tested with three separate subsets of the data. One subset is used for training the network, one for validation, and one for testing the trained networks. The multi-platform data set contains 53 targets and 18 of these targets were chosen at random for the training subset from the remaining 35 targets, 10 were chosen at random for the validation subset and the remaining 25 were used for the testing subset. In order to make the networks robust to target location variation within the ROI, multiple ROI's over the target were used for the training, validation, and testing subsets where the target was contained in different locations within the ROI. For every target and background sample in the three co-registered sonar images, we used 4 ROI's hence resulting a training set consisting of 72 sets of three target and 72 sets of three background ROI's chosen from this multi-platform sonar data set. The validation and testing subset were formed similarly yielding 40 sets of three target and 40 sets of three background ROI's for the validation set and 100 sets of three target and 100 sets of three background ROI's for the testing set and both were extracted from the three co-registered sonar images. The decisions from the two CCA-based local decision makers are then computed for the training subset and are used to train the BPNN. For the training of the BPNN, five random weight initializations were used

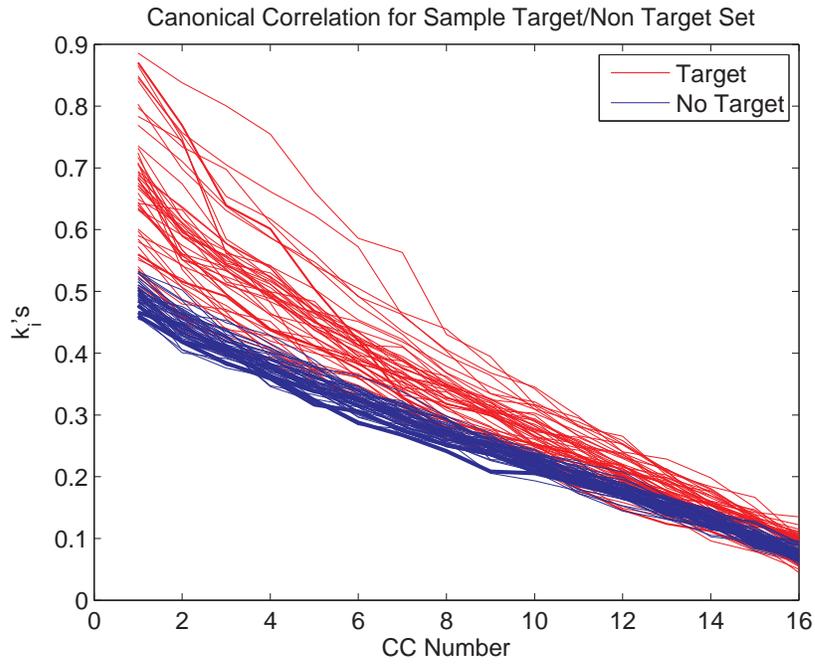
and the BPNN which performed the best on the validation set was selected for testing.

For the training of the PNN again three subsets of the data were used for training, validation and testing. The input to the PNN is the observation from the environment,  $\mathbf{z}_0$ . The observation was chosen to be a set of statistical attributes that are taken from the three processed ROI's. These attributes are the mean, variance, and skew of data within each of the ROI's from the set of three co-registered sonar images. Thus,  $\mathbf{z}_0$  is a 9-D observation vector provided to the fusion center. The subset for the training consisted of 72 target and 72 background observations. Similarly, the validation consisted of 40 target and 40 background observations and the testing consisted of 100 target and 100 background observations. Since the weights in the pattern layer are fixed for the PNN, 10 different initializations of the variance of the Gaussian were used for the training of the PNN. The network which performed the best on the validation set was selected for testing. The resulting trained PNN is relatively small due the binary hypothesis problem as the PNN consists of only two pools of neurons in the pattern layer.

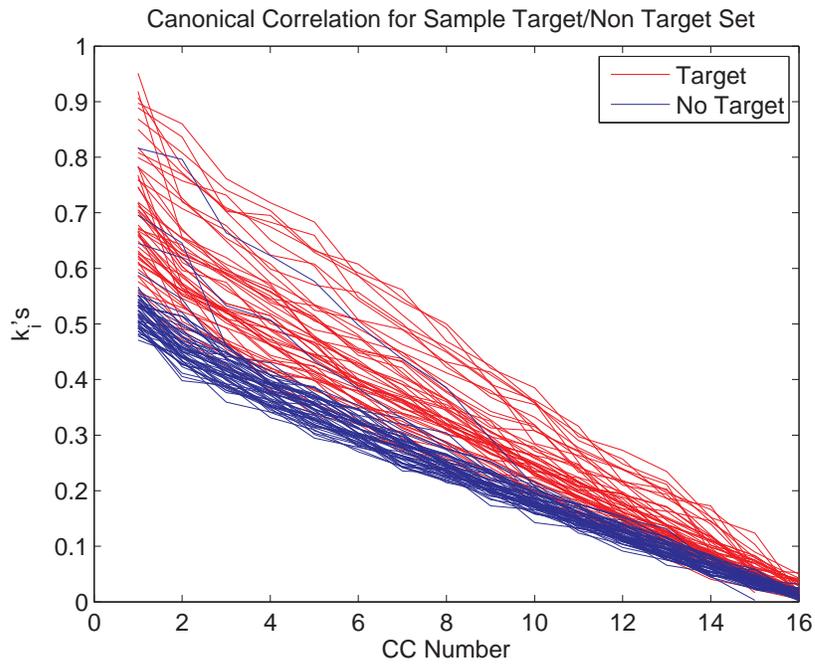
### 5.3.2 Results and Observations

To show the separability of the principal canonical correlations for ROI's that contain targets over background and those that contain only background, a test was conducted on the entire target set ROI's and a same size randomly selected set of ROI's containing only background clutter for two local detectors. The plots of 16 canonical correlations of ROI's containing targets and those containing background only are shown in Figures 5.3(a) and (b) for the HF-BB<sub>1</sub> and BB<sub>1</sub>-BB<sub>2</sub> detectors, respectively. As can be seen, for both pairs of sonar images there is suitable separation among the canonical correlations formed from target ROI's and those from background ROI's, especially for the principal canonical correlations.

Figures 5.4(a) and (b) show the histograms of the log-likelihood ratio values of



(a) High Frequency Broadband Detector



(b) Broadband Broadband Detector

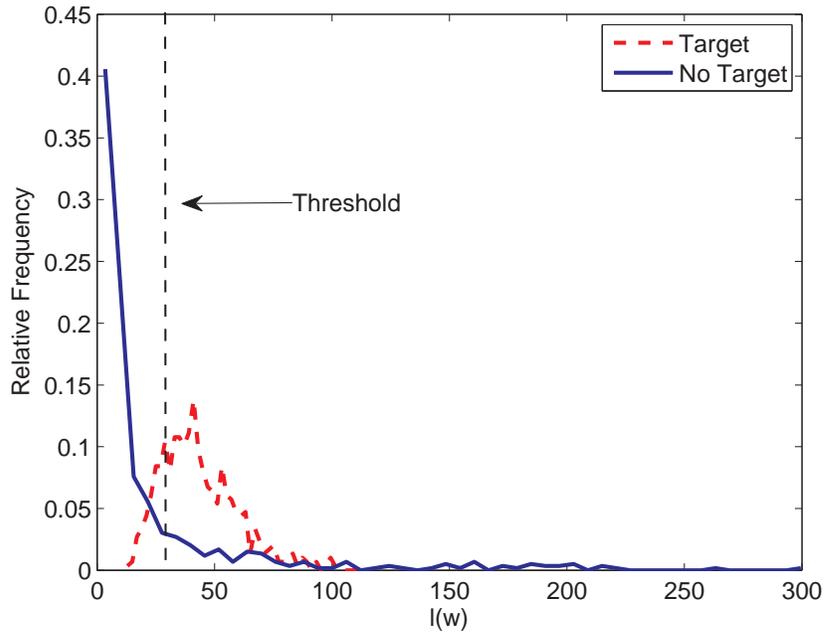
**Figure 5.3:** Plot of Canonical Correlations for Target and Background for both Detectors.

one target ROI and one random background ROI in the HF-BB<sub>1</sub> and BB<sub>1</sub>-BB<sub>2</sub> cases, respectively. Using the entire set of target and background ROI's, the optimum thresholds for the local decision makers was determined to be 26.4 for the HF-BB<sub>1</sub> and 23.9 for the BB<sub>1</sub>-BB<sub>2</sub>, respectively. These thresholds are shown by dotted vertical lines in Figures 5.4(a) and (b). These values were then used as detection thresholds for all the ROI's. That is, any block within an ROI pair whose log-likelihood ratio falls above the relevant threshold is flagged as a target block in either HF-BB<sub>1</sub> or BB<sub>1</sub>-BB<sub>2</sub> pairs and the decision is set to the fusion center.

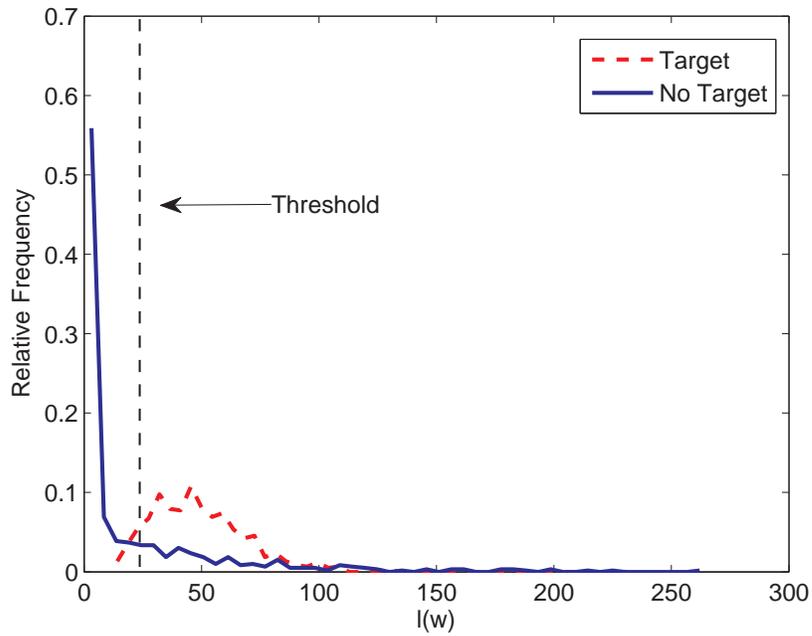
In order to determine the threshold for the fusion rule a "1-0" cost function is used, i.e  $C_{10} = C_{01} = 1$  and  $C_{11} = C_{00} = 0$ . Additionally, for the prior probabilities,  $P_0$  and  $P_1$ , are considered to be equal, i.e.  $P_0 = P_1 = 1/2$ . Therefore, the threshold for the fusion rule in (5.4) is 1.

The system is then implemented on the entire NSWC multi-platform sonar imagery data. The individual coherence-based local decision makers performed marginally well. More specifically, the HF-BB<sub>1</sub> detector detected 51 of the 53 targets with an average of 10.13 false detections per image. While the BB<sub>1</sub>-BB<sub>2</sub> detector performed a little worse with detection of 49 of the 53 targets and an average of 9.86 false detections per image. However, when the decisions of these two detectors are fused, the performance was greatly improved. After the fusion process, the system detected 53 of the 53 targets with an average of only 7.1 false detections per image. The ROC curves for the individual detectors and the fusion of the two are presented in Figure 5.5 and were generated from the entire multi-platform sonar data set. At the knee point of the ROC for the HF-BB<sub>1</sub> detector we have  $P_d = 95\%$  and  $P_{fa} = 5\%$ . While for the BB<sub>1</sub>-BB<sub>2</sub> detector the knee point exhibits  $P_d = 92\%$  and  $P_{fa} = 8\%$ . After the fusion of the two detectors, the knee point of the fused ROC gives  $P_d = 99\%$  and  $P_{fa} = 1\%$ , which demonstrates excellent overall detection performance of the proposed system.

After fusion all of the targets were detected and the number of false alarms per

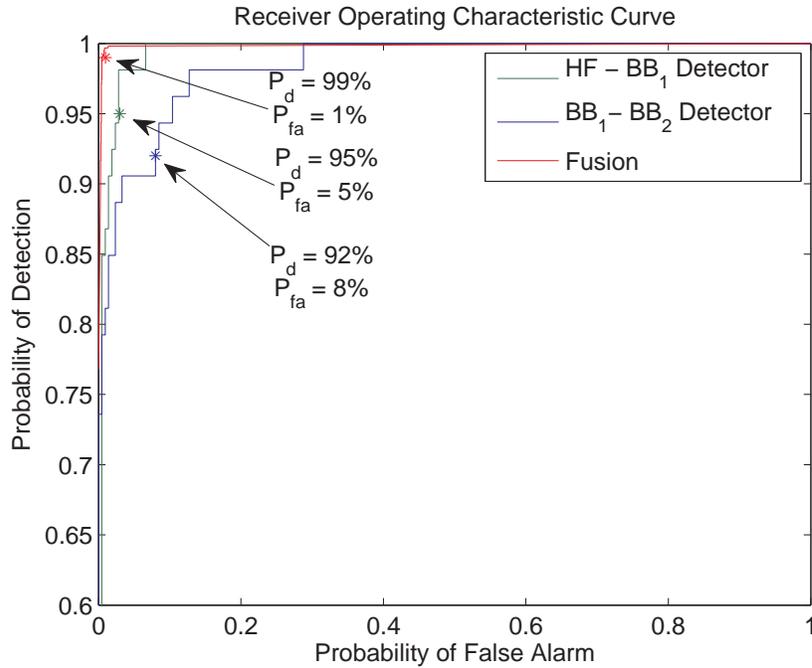


(a) High Frequency Broadband Detector



(b) Broadband Broadband Detector

**Figure 5.4:** Histogram of Example Log-Likelihood Values for Target and Background for both Detectors.



**Figure 5.5:** ROC Curves for the Individual Detectors and the Fusion

image was greatly reduced. The targets that were missed by the local decision makers were at very close range and had faint signatures in all the three images. This is due to the fact that these targets were close to the tract of the AUV. Consequently, the signature was extremely dark and hard to visually discern in all three images hence leading to overall low coherence and subsequent misdetection. But after the fusion rule even the faint target signatures were detected and this can be attributed to the fusion center having it's own observation of the environment. Overall, the distributed detection system performed extremely well given the small number of targets and non-targets used to form the detection threshold and showed vast improvement over the single sensor detector in Section 4.3 as well as the dual disparate detector in Section 4.4.

## 5.4 Conclusion

In this chapter, a distributed detection scheme is developed where several dual-platform detectors are used to generate local decisions. The fusion center receives these local decisions together with its own observation to generate a final decision. The developed fusion rule is based on the work in [48]. The incorporation of the observation vector in the fusion rule helps to greatly increase the probability of detection and decrease the false alarm rate. By using multiple dual-platform detectors one is able to take advantage of more than two observations from the environment hence providing a higher confidence detection decision at the fusion center.

When the developed distributed detection scheme was applied to the multi-platform data set provided by the NSWC-Panama City it was shown that the fusion process significantly improved the probability of correct detection while reducing the incidence of a false alarm when compared with the results of the single dual disparate detector in Chapter 4. All targets were detected with an average of only 7 false detections per image on the entire multi-platform sonar data set. When comparing the results of the individual dual disparate detectors to the results of the fusion, there was a 4% improvement in the  $P_d$  over the individual HF-BB<sub>1</sub> detector and a 7% improvement in the  $P_d$  over the individual BB<sub>1</sub>-BB<sub>2</sub> detector and a substantial reduction in the number of false alarms over either of the individual detectors. Clearly, this indicates that the use of multiple sonar types improves the detection performance by better capturing the target characteristics. Moreover, the use of multiple decisions from the environment in a realistic NSA scenario can effectively implement multi-platform detection by reducing the amount of information that needs to be sent to the fusion center while offering high probability of detection and low false alarm rate.

In the next chapter we investigate the effect of sample support on the Gauss-Gauss detector performance. More specifically, two cases are examined, the sample rich and sample poor cases. In the sample rich case, the number of samples is greater

than the data vector dimension while for the sample poor case the opposite is true. We develop a method for finding the eigenvalues and eigenvectors of the signal-to-noise ratio matrix for both cases as well as the associated log-likelihood function and J-divergence criterion. The results in the next chapter are important as they provide some guidelines on how to successfully design and implement coherence-based detectors using limited data samples.

# CHAPTER 6

## EFFECT OF SAMPLE SUPPORT ON LINEAR AND KERNEL GAUSS-GAUSS DETECTION

### 6.1 Introduction

Detection of a Gaussian signal in Gaussian noise (Gauss-Gauss detection) was addressed in Chapter 3.2.2 where it was shown how the log-likelihood ratio can be formulated in terms of the eigenvalues and eigenvectors of the “signal-to-noise ratio” matrix which involves the signal covariance matrix and the inverse of the noise covariance matrix. Moreover, a low-rank Gauss-Gauss detector can be constructed based upon the J-divergence which is a measure of distance between the two hypotheses.

Computing the signal-to-noise ratio matrix requires the knowledge of data covariance matrices corresponding to the two hypotheses, which in practice have to be estimated from a limited number of samples drawn from the corresponding observations. However, little attention has been paid to the study of the effects of sample support on the eigenvalues of the signal-to-noise ratio matrix. More specifically, we would like to determine just how poor can sample support become before the eigenvalues of the signal-to-noise ratio matrix cease to carry useful information and how does this affect the validity of the J-divergence and log-likelihood function. Thus, in this chapter, we consider Gauss-Gauss detection when data covariance matrices under the two hypotheses are estimated from a limited number of data samples. We study two sample support scenarios namely, the *sample rich* and *sample poor* cases. In the sample rich case, the number of samples is greater than the data vector dimension while for the sample poor case the opposite is true. We then develop a method for

finding the eigenvalues and eigenvectors of the signal-to-noise ratio matrix for both cases as well as the associated log-likelihood function and J-divergence criterion. We show, through a numerical example, that as the number of samples becomes closer to or less than the dimension of the data vector, the eigenvalues of the signal-to-noise ratio matrix become defective leading to poor detection results and a superficial J-divergence measure for detectability. Although the sample poor scenarios may not occur frequently for the linear detector, they do indeed happen in the kernel nonlinear version of this detector. This is due to the fact that the dimension of the mapped data in the high dimensional feature space is typically much higher than the sample support (e.g. Gaussian kernel).

The idea behind kernel Gauss-Gauss detection is to use a nonlinear mapping to map the input space into a high-dimensional feature space, in which the cases are linearly separable. Perhaps the most intriguing aspect is that the high dimensional nonlinear mappings are never explicitly computed and all computations are carried out in the original low dimensional space using the *kernel trick* [36] - [38]. Kernel-based detection has attracted quite a bit of attention lately. In [56], a kernel matched filter detector was developed by extending the linear matched filter. The method in [57], developed the kernel version of the well-known RX anomaly detector for hyperspectral imagery. The kernel matched subspace detector was developed in [58] by formulating the generalized log-likelihood ratio test in the high-dimensional feature space. In this chapter, we also extend the sampled version of the linear Gauss-Gauss detector to the kernel version by formulating the log-likelihood in the high-dimensional feature space using the kernel trick.

The outline of this chapter is as follows. Section 6.2 develops the empirical version of the Gauss-Gauss detector where both the sample rich and sample poor cases are explored. In Section 6.3 the non-linear kernel Gauss-Gauss detector is extended. Section 6.4 presents a numerical example with the corresponding eigenvalue,

J-divergence, and likelihood ratio analysis for the linear sample rich and sample poor cases. Finally, conclusions are made in Section 6.5.

## 6.2 Gauss-Gauss Detection and Data Sample Support

Let us, for the sake of completeness, review the classical detection problem (see Chapter 3.2.1), which involves deciding between two hypotheses, the one of noise only and the other of signal plus noise. Assume we have an observation  $\mathbf{x} \in \mathbb{R}^N$  which is a normal random vector with zero mean and covariance matrix  $R$ . We wish to test the hypothesis  $H_0 : R = R_0$  i.e. noise alone versus  $H_1 : R = R_1$  i.e signal plus noise,  $R_1 = R_0 + R_s$  where  $R_0$  is the covariance matrix of the noise alone and  $R_s$  is the covariance matrix of the signal. We can also view this as a two channel problem [27] where  $\mathbf{x} = \mathbf{n}$  is noise alone with covariance  $R_0$  and  $\mathbf{y} = \mathbf{s} + \mathbf{n}$  is signal plus noise with covariance  $R_1$  under hypothesis  $H_1$ .

The log-likelihood ratio test (LRT) that minimizes the risk involved in deciding between  $H_0$  and  $H_1$  leads to

$$\gamma(\mathbf{x}) = \begin{cases} 1 \sim H_1, & \text{when } l(\mathbf{x}) > \lambda \\ 0 \sim H_0, & \text{with } l(\mathbf{x}) \leq \lambda \end{cases}$$

where  $l(\mathbf{x}) = \mathbf{x}^H \mathbf{Q} \mathbf{x}$  is the log-likelihood ratio and  $\mathbf{Q} = R_0^{-1} - R_1^{-1}$  and matrix  $\mathbf{Q}$  is given by

$$\mathbf{Q} = R_0^{-H/2} (\mathbf{I} - \mathbf{S}^{-1}) R_0^{-1/2}, \quad (6.1)$$

where the “signal-to-noise ratio” matrix [35],  $\mathbf{S}$ , is defined as

$$\mathbf{S} = R_0^{-1/2} R_1 R_0^{-H/2} \quad (6.2)$$

The J-divergence [35] between the two hypotheses, which is a global measure of

the separability or detectability of the two hypotheses is

$$J = E_{H_1}[l(\mathbf{x})] - E_{H_0}[l(\mathbf{x})] \quad (6.3)$$

$$= \text{tr}(\mathbf{S} + \mathbf{S}^{-1} - 2\mathbf{I}). \quad (6.4)$$

The matrix  $\mathbf{S}$  has an eigenvalue decomposition as,

$$\mathbf{S} = R_0^{-1/2} R_1 R_0^{-H/2} = \mathbf{U} \Lambda \mathbf{U}^H. \quad (6.5)$$

The J-divergence in (6.4) can be rewritten as

$$\begin{aligned} J &= \text{tr}(\Lambda + \Lambda^{-1} - 2\mathbf{I}) \\ &= \sum_{i=1}^N (\lambda_i + \lambda_i^{-1} - 2). \end{aligned} \quad (6.6)$$

Moreover, the log-likelihood function can also be expressed in terms of eigenvalues and eigenvectors of  $\mathbf{S}$  as,

$$l(\boldsymbol{\xi}) = \boldsymbol{\xi}^H \mathbf{U} (\mathbf{I} - \Lambda^{-1}) \mathbf{U}^H \boldsymbol{\xi} \quad (6.7)$$

where  $\boldsymbol{\xi} = R_0^{-1/2} \mathbf{x}$ .

Thus, we only need to solve the eigenvalue problem in (6.5) to obtain the eigenvalues  $\lambda_i$  and eigenvectors of  $\mathbf{S}$  for computing the J-divergence and log-likelihood function in (6.6) and (6.7), respectively.

In practice, however, the covariance matrices are estimated from a limited number of data samples. Consider the sample data matrices  $X = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M] \in \mathbb{R}^{m \times M}$  and  $Y = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_M] \in \mathbb{R}^{m \times M}$  with  $m$  being the dimensions of  $\mathbf{x}$  and  $\mathbf{y}$  and  $M$  being the number of samples (typically  $m < M$ ). Here  $\mathbf{x}_k$  and  $\mathbf{y}_k$  are independent realizations of the random vectors  $\mathbf{x}$  and  $\mathbf{y}$ , defined before. Note that for the following derivations ‘ $\hat{\cdot}$ ’ implies the sampled version of the corresponding matrix. Thus, the sample data covariance matrix associated with  $\mathbf{x}$  is  $\hat{R} = \mathbf{X} \mathbf{X}^H = \sum_{k=1}^M \mathbf{x}_k \mathbf{x}_k^H$ . To test the hypothesis  $H_0$  versus  $H_1$  the sample covariances are used in place of the theoretical ones. We now have two cases to consider, the *sample rich case*, i.e.  $m \leq M$  and the *sample poor case*, i.e.  $m > M$ .

### 6.2.1 Sample Rich Case ( $m < M$ )

The sampled version of  $Q$  is

$$\begin{aligned}\hat{Q} &= (XX^H)^{-1} - (YY^H)^{-1} \\ &= (XX^H)^{-H/2}(\mathbf{I} - \hat{S}^{-1})(XX^H)^{-1/2},\end{aligned}\tag{6.8}$$

where the sampled version of  $S$  is defined as

$$\hat{S} = (XX^H)^{-1/2}(YY^H)(XX^H)^{-H/2}.\tag{6.9}$$

Since  $l(\mathbf{x})$  is scalar using the cyclic property of the trace we can write,

$$\begin{aligned}l(\mathbf{x}) &= \text{tr}(\mathbf{x}^H \hat{Q} \mathbf{x}) \\ &= \text{tr}(\mathbf{x} \mathbf{x}^H (XX^H)^{-1}) - \text{tr}(\mathbf{x} \mathbf{x}^H (YY^H)^{-1}).\end{aligned}\tag{6.10}$$

Thus, using (6.3) we can write the J-divergence as

$$J = \text{tr}[(YY^H)(XX^H)^{-1} + (XX^H)(YY^H)^{-1} - 2\mathbf{I}],\tag{6.11}$$

where  $E_{H_0}[\mathbf{x} \mathbf{x}^H] = \hat{R}_0 = XX^H$  and  $E_{H_1}[\mathbf{x} \mathbf{x}^H] = \hat{R}_1 = YY^H$ . If we then rewrite  $(XX^H) = (XX^H)^{1/2}(XX^H)^{H/2}$ , then we get

$$\begin{aligned}J &= \text{tr}[(YY^H)(XX^H)^{-H/2}(XX^H)^{-1/2} \\ &\quad + (XX^H)^{1/2}(XX^H)^{H/2}(YY^H)^{-1} - 2\mathbf{I}]\end{aligned}\tag{6.12}$$

or

$$J = \text{tr}(\hat{S} + \hat{S}^{-1} - 2\mathbf{I}).\tag{6.13}$$

which is the sampled version of (6.4).

Matrix  $\hat{S}$  has an orthogonal decomposition

$$\hat{S} = (XX^H)^{-1/2}(YY^H)(XX^H)^{-H/2} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H\tag{6.14}$$

where  $\mathbf{\Lambda}$  and  $\mathbf{U}$  can be solved using the sampled generalized eigenvalue problem

$$(XX^H)^{-1/2}(YY^H)(XX^H)^{-H/2}\mathbf{U} = \mathbf{U}\mathbf{\Lambda}.\tag{6.15}$$

which together with (6.13) reconfirms (6.6) for this sample rich case.

Substituting the orthogonal decomposition of  $\hat{S}$  into (6.8) we can rewrite  $\hat{Q}$  as

$$\hat{Q} = (XX^H)^{-H/2}U(I - \Lambda^{-1})U^H(XX^H)^{-1/2} \quad (6.16)$$

If we define  $D \triangleq (XX^H)^{-H/2}U$ , then (6.16) can be rewritten as

$$\hat{Q} = D(I - \Lambda^{-1})D^H \quad (6.17)$$

And, the log-likelihood function becomes

$$l(\mathbf{z}) = \mathbf{z}^H D(I - \Lambda^{-1})D^H \mathbf{z} \quad (6.18)$$

Using the definition of  $D$ , (6.15) can be converted to the following generalized eigenvalue problem

$$\begin{aligned} (XX^H)^{-H/2}(XX^H)^{-1/2}(YY^H)(XX^H)^{-H/2}U &= (XX^H)^{-H/2}U\Lambda \\ (YY^H)D &= (XX^H)D\Lambda, \end{aligned} \quad (6.19)$$

Therefore, matrices  $D$  and  $\Lambda$  can be found from this generalized eigenvalue problem which can then be used in (6.18) to compute the log-likelihood and (6.13) to compute the J-divergence.

### 6.2.2 Sample Poor Case ( $m > M$ )

If we assume that  $m > M$ , i.e. sample poor case, and that  $\text{rank}(XX^H) = p \leq M$ . In the linear case  $p = M$  if the samples are drawn independently. The case  $p \leq M$  is important when the sampled data is mapped using a nonlinear mapping for the kernel detector and the  $\text{rank}(XX^H)$  is not guaranteed to be  $M$ . Using the eigenvalue decomposition of  $X$  and  $Y$  we have

$$\begin{aligned} X &= V_1\Lambda_X V_2^H \\ Y &= W_1\Lambda_Y W_2^H, \end{aligned} \quad (6.20)$$

where  $V_1$  and  $V_2$  are the left and right transformation matrices for  $X$  and similarly  $W_1$  and  $W_2$  are the left and right transformation matrices for  $Y$ , and

$$\Lambda_X = \begin{bmatrix} \Lambda_X(p) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \Lambda_Y = \begin{bmatrix} \Lambda_Y(p) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (6.21)$$

where  $\Lambda_X(p)$  and  $\Lambda_Y(p)$  are  $p \times p$  diagonal matrices with diagonal elements that are eigenvalues of  $X$  and  $Y$  data matrices, respectively.

We can then rewrite (6.9) as,

$$\hat{S} = (V_1 \Lambda_X^2 V_2^H)^{-1/2} (W_1 \Lambda_Y^2 W_2^H) (V_1 \Lambda_X^2 W_2^H)^{-H/2}. \quad (6.22)$$

where all matrix inversion operations imply the use of the Moore-Penrose pseudo-inverse [51] for the corresponding matrix. Therefore,  $\hat{S}$  will be of rank  $p$  and hence we can rewrite it as,

$$\begin{aligned} \hat{S}_p &= (V_{1,p} \Lambda_X^2(p) V_{1,p}^H)^{-1/2} (W_{1,p} \Lambda_Y^2(p) W_{2,p}^H) (V_{1,p} \Lambda_X^2(p) V_{1,p}^H)^{-H/2} \\ &= U_p \Lambda(p) U_p^H \end{aligned} \quad (6.23)$$

where  $X$  and  $Y$  are expressed as  $X = V_{1,p} \Lambda_X(p) V_{2,p}^H$  and  $Y = W_{1,p} \Lambda_Y(p) W_{2,p}^H$  where  $V_{1,p}$  and  $V_{2,p}$  contain the first  $p$  columns of  $V_1$  and  $V_2$  corresponding to the non-zero eigenvalues of  $X$ , similarly for  $Y$ ,  $W_{1,p}$  and  $W_{2,p}$  contain the first  $p$  columns of  $W_1$  and  $W_2$  corresponding to the non-zero eigenvalues of  $Y$  and

$$\Lambda = \begin{bmatrix} \Lambda(p) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (6.24)$$

and  $U_p$ , which is of dimension  $m \times p$ , contains the first  $p$  columns of  $U$  corresponding to the non-zero eigenvalues of  $\hat{S}$ .

If we define  $D_p \triangleq (X X^H)^{-H/2} U_p$ , then  $\hat{S}_p$  can be converted to a generalized eigenvalue problem as in (6.19), but in the reduced rank subspace,

$$(Y Y^H) D_p = (X X^H) D_p \Lambda(p). \quad (6.25)$$

From  $D_p \triangleq (XX^H)^{-H/2}U_p$  it is obvious that  $D_p$  is in the span of  $X$ . Thus, we can write  $D_p = XA_p$  yielding

$$(YY^H)XA_p = (XX^H)XA_p\Lambda(p). \quad (6.26)$$

Pre-multiplying both sides of (6.26) by  $X^H$  gives

$$(X^HY)(Y^HX)A_p = (X^HX)(X^HX)A_p\Lambda(p), \quad (6.27)$$

which can be solved to give  $\Lambda(p)$  and  $A_p$ . For the sample poor case, the J-divergence can be found from the eigenvalues of  $\hat{S}_p$  using (6.13) for only  $p$  non-zero eigenvalues.

Now, using the definitions of  $D_p$  and  $\Lambda(p)$  we can rewrite (6.17) as

$$\hat{Q} = XA_p(I - \Lambda^{-1}(p))A_p^HX^H \quad (6.28)$$

and the log-likelihood in (6.18) can be expressed as,

$$l(\mathbf{x}) = \mathbf{x}^HX(I - A_p\Lambda^{-1}(p)A_p^H)X^H\mathbf{x}. \quad (6.29)$$

That is, the log-likelihood can be found using  $\Lambda(p)$  and  $A_p$  which are obtained by solving the generalized eigenvalue problem in (6.27). Note that the formulations in (6.27) and (6.29) are very relevant to the development in the next section as the terms  $(X^HX)$ ,  $(XY^H)$ ,  $(Y^HX)$ , and  $(\mathbf{x}^HX)$  are already in dot product form and further due to the high dimensionality of the feature space the sample poor conditions are inevitable.

### 6.3 Kernel Gauss-Gauss Detection

In this section, based on our findings about the sample support of the log-likelihood and J-divergence we develop the kernel version of the Gauss-Gauss detector. The motivation behind the kernel detector is that by using a nonlinear mapping we can map the data in the input space into a higher dimensional space where the features are linearly separable. From the previous formulations the detector is already in terms of

the dot products of the data matrices and the nonlinear mapping can be carried out in the original input space using the *kernel trick* [36]. The inner products are then replaced by kernel functions that can be computed in the input space that satisfy Mercer's conditions.

We can define the mapped data matrices  $\Phi = [\phi(\mathbf{x}_1), \phi(\mathbf{x}_2), \dots, \phi(\mathbf{x}_M)] \in \mathbb{R}^{m' \times M}$  and  $\Psi = [\phi(\mathbf{y}_1), \phi(\mathbf{y}_2), \dots, \phi(\mathbf{y}_M)] \in \mathbb{R}^{m' \times M}$ , where typically  $m' \gg m$ . Here each pair of samples  $\mathbf{x}_i$  and  $\mathbf{y}_i$  are mapped using high-dimensional kernel producing [36] - [38] mapping function  $\phi(\cdot)$ .

We wish to test the hypothesis  $H_0 : R_\phi = R_{\phi_0}$  versus  $H_1 : R_\phi = R_{\phi_1}$  in the implicit high-dimensional feature space using the kernel trick [36]. The log-likelihood ratio test that decides between  $H_0$  and  $H_1$  leads to

$$\gamma(\phi) = \begin{cases} 1 \sim H_0, & \text{when } l(\phi) > \lambda \\ 0 \sim H_1, & \text{when } l(\phi) \leq \lambda \end{cases}$$

where  $l(\phi) = \phi^H(\mathbf{x})Q_\phi\phi(\mathbf{x})$  and  $Q_\phi = R_{\phi_0}^{-1} - R_{\phi_1}^{-1}$ . Here it is assumed that the data in the mapped high-dimensional feature space is Gaussian. Using the previous formulations (6.28) and (6.29) in Section 6.2.2, and the mapped data matrices  $\Phi$  and  $\Psi$ , we can write  $Q_\phi$  as,

$$Q_\phi = \Phi A_\phi (\mathbf{I} - \Lambda_\phi^{-1}) A_\phi^H \Phi^H \quad (6.30)$$

and hence the log-likelihood can be expressed as

$$l(\phi(\mathbf{x})) = \mathbf{k}_{\phi\Phi}^H (\mathbf{I} - A_\phi \Lambda_\phi^{-1} A_\phi^H) \mathbf{k}_{\phi\Phi}. \quad (6.31)$$

where  $\mathbf{k}_{\phi\Phi} = \phi^H(\mathbf{x})\Phi = [k(\mathbf{x}, \mathbf{x}_1), \dots, k(\mathbf{x}, \mathbf{x}_M)]^H$  and  $k(\mathbf{x}, \mathbf{x}_i) = \phi^H(\mathbf{x})\phi(\mathbf{x}_i)$  is a kernel function that satisfies Mercer's conditions [36]. As in (6.27), matrices  $A_\phi$  and  $\Lambda_\phi$  can be found by solving the generalized eigenvalue problem

$$K_{\Phi\Psi} K_{\Phi\Psi}^H A_\phi = K_{\Phi\Phi} K_{\Phi\Phi} A_\phi \Lambda_\phi \quad (6.32)$$

where  $K_{\Phi\Phi} = \Phi^H \Phi = [k(\mathbf{x}_i, \mathbf{x}_j)]_{ij}$  and  $K_{\Phi\Psi} = \Phi^H \Psi = [k(\mathbf{x}_i, \mathbf{y}_j)]_{ij}$  are the Gram kernel matrices. Upon solving the generalized eigenvalue problem in (6.32), one can

then find the J-divergence using (6.13) and the log-likelihood using (6.31), which is also expressed only in terms of kernels.

Note that here we assumed that the data in the high-dimensional feature space is centered (zero-meaned). However, this cannot be guaranteed in this implicit and possibly unknown space. To implement this centering using the kernel trick let the mean of the mapped data,  $\bar{\phi}$  be defined as  $\bar{\phi} = (1/M) \sum_{m=1}^M \phi(\mathbf{x}_m)$ . Then, we can define the ‘‘centered’’ version of  $\phi(\mathbf{x})$  as  $\tilde{\phi}(\mathbf{x}) = \phi(\mathbf{x}) - \bar{\phi}$ . We now can form the centered version of  $K_{\Phi\Phi}$  as  $\tilde{K}_{\Phi\Phi} = [\tilde{k}(\mathbf{x}_i, \mathbf{x}_j)]_{ij}$ , where

$$\begin{aligned} \tilde{k}(\mathbf{x}_i, \mathbf{x}_j) &= \tilde{\phi}^H(\mathbf{x}_i)\tilde{\phi}(\mathbf{x}_j) \\ &= k(\mathbf{x}_i, \mathbf{x}_j) - \frac{1}{M} \sum_{m=1}^M k(\mathbf{x}_i, \mathbf{x}_m) - \frac{1}{M} \sum_{n=1}^M k(\mathbf{x}_n, \mathbf{x}_j) + \frac{1}{M^2} \sum_{m=1}^M \sum_{n=1}^M k(\mathbf{x}_m, \mathbf{x}_n) \\ &= k(\mathbf{x}_i, \mathbf{x}_j) - \frac{1}{M} \mathbf{1}^T \mathbf{k}_{\phi\Phi} - \frac{1}{M} \mathbf{1}^T \mathbf{k}_{\Phi\phi} + \frac{1}{M^2} \mathbf{1}^T K_{\Phi\Phi} \mathbf{1} \end{aligned} \quad (6.33)$$

where  $\mathbf{k}_{\phi\Phi}$  is defined as before and  $\mathbf{1} = [1, \dots, 1]^T$ . Now using (6.33) for every element of  $\tilde{K}_{\Phi\Phi}$  yields

$$\begin{aligned} \tilde{K}_{\Phi\Phi} &= K_{\Phi\Phi} - \frac{1}{M} \mathbf{1}\mathbf{1}^T K_{\Phi\Phi} - \frac{1}{M} K_{\Phi\Phi} \mathbf{1}\mathbf{1}^T + \frac{1}{M^2} \mathbf{1}\mathbf{1}^T K_{\Phi\Phi} \mathbf{1}\mathbf{1}^T \\ &= P_1^\perp K_{\Phi\Phi} P_1^\perp \end{aligned} \quad (6.34)$$

where  $P_1^\perp = \mathbf{I} - \frac{1}{M} \mathbf{1}\mathbf{1}^T$  is a centering matrix [36].

Using (6.34) in (6.32) yields the ‘‘centered generalized eigenvalue’’ problem as

$$P_1^\perp K_{\Phi\Psi} P_1^\perp K_{\Phi\Psi}^H P_1^\perp A_\phi = P_1^\perp K_{\Phi\Phi} P_1^\perp K_{\Phi\Phi} P_1^\perp A_\phi \Lambda_\phi. \quad (6.35)$$

To compute the log-likelihood function in the kernel domain the centered version of  $\mathbf{k}_{\phi\Phi}$  is also needed. This can be obtained using (6.33) as

$$\begin{aligned} \tilde{\mathbf{k}}_{\phi\Phi} &= \mathbf{k}_{\phi\Phi} - \frac{1}{M} \mathbf{1}\mathbf{1}^T \mathbf{k}_{\phi\Phi} - \frac{1}{M} K_{\Phi\Phi} \mathbf{1} + \frac{1}{M^2} \mathbf{1}\mathbf{1}^T K_{\Phi\Phi} \mathbf{1} \\ &= P_1^\perp (\mathbf{k}_{\phi\Phi} - \frac{1}{M} K_{\Phi\Phi} \mathbf{1}) \end{aligned} \quad (6.36)$$

Using (6.36) in (6.31) gives the ‘‘centered log-likelihood’’ as

$$l(\phi) = (\mathbf{k}_{\phi\Phi} - \frac{1}{M} K_{\Phi\Phi} \mathbf{1})^H P_1^\perp (\mathbf{I} - A_\phi \Lambda_\phi^{-1} A_\phi^H) P_1^\perp (\mathbf{k}_{\phi\Phi} - \frac{1}{M} K_{\Phi\Phi} \mathbf{1}). \quad (6.37)$$

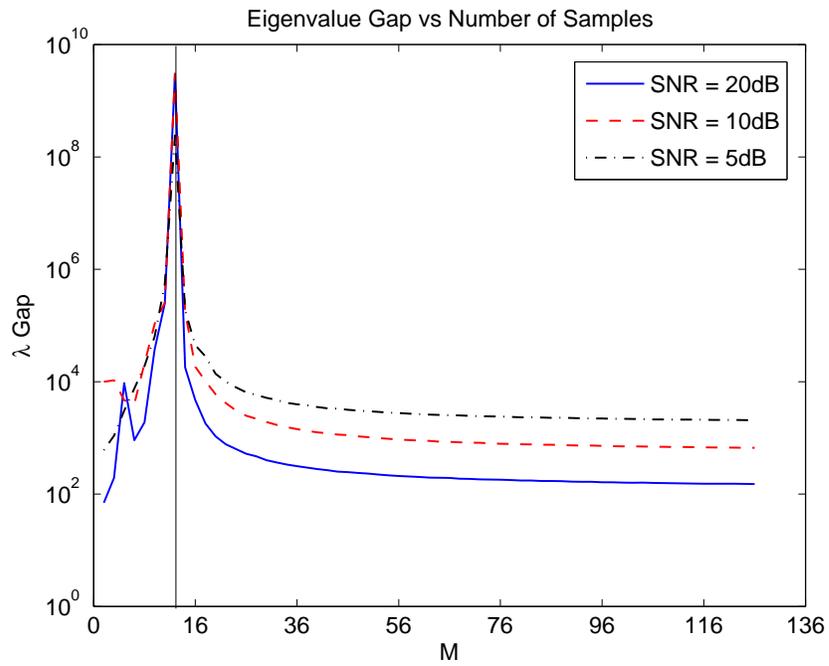
Therefore, both the log-likelihood and the generalized eigenvalue problem can be written in terms of the centered versions of the Gram kernel matrices and the kernel vector hence guaranteeing that our data will be centered in the feature space.

## 6.4 Results and Observations

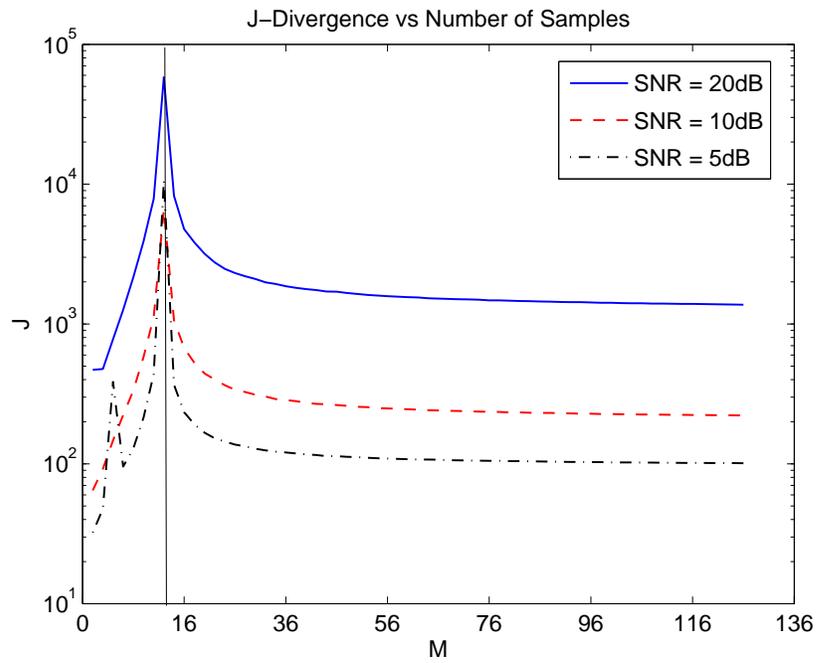
An experiment is conducted to show the effectiveness of the detector for the sample rich and the sample poor cases. Under  $H_0$ ,  $\mathbf{x} = \mathbf{n}$  is a zero-mean white Gaussian vector process. The signal,  $\mathbf{s}$ , is formed from an arbitrary matrix  $B$  and another zero-mean white Gaussian noise vector,  $\boldsymbol{\eta}$  with  $R_\eta = 1.25\mathbf{I}$ , such that  $\mathbf{s} = B\boldsymbol{\eta}$ . Therefore, under  $H_1$  we have  $\mathbf{y} = B\boldsymbol{\eta} + \mathbf{n}$ . The sample data matrices are formed for the two cases each with a dimension of  $m = 12$  and having a total of 256 samples. For the sample rich and sample poor cases three SNR scenarios of  $20dB$ ,  $10dB$ , and  $5dB$  were considered. The matrix  $\hat{S}$  is formed to build the log-likelihood function and compute the J-divergence. The theoretical (for large sample support) J-divergence for the three SNR cases are found to be 1371, 218, and 100 respectively.

The eigenvalues of  $\hat{S}$  for the sample rich and sample poor cases are computed from (6.19) and (6.27), respectively. The J-divergence for both cases is computed using (6.13). The number of samples,  $M$  was varied from 256 to 12 for the sample rich case and from 4 to 11 for the sample poor case. For each case the eigenvalues are found, J-divergence is computed and a separate test set of 256 samples is then applied to the detector. The experiment was repeated for 1000 Monte Carlo trials, where for each trial the signal model is used to generate the data, but for a fixed  $B$  matrix.

The plots of the eigenvalue gap between the largest and smallest eigenvalues ( $\lambda_1/\lambda_{12}$ ), J-divergence, probability of detection ( $P_d$ ), and probability of false alarm ( $P_{fa}$ ), versus the number of samples are presented in Figures 6.1-6.4, respectively. These plots represent the average results for 1000 Monte Carlo trials. In each figure, three plots for the SNR cases of 20, 10, and  $5dB$  are presented. The vertical line



**Figure 6.1:** Eigenvalue Gap versus Number of Samples.



**Figure 6.2:** J-divergence versus Number of Samples.

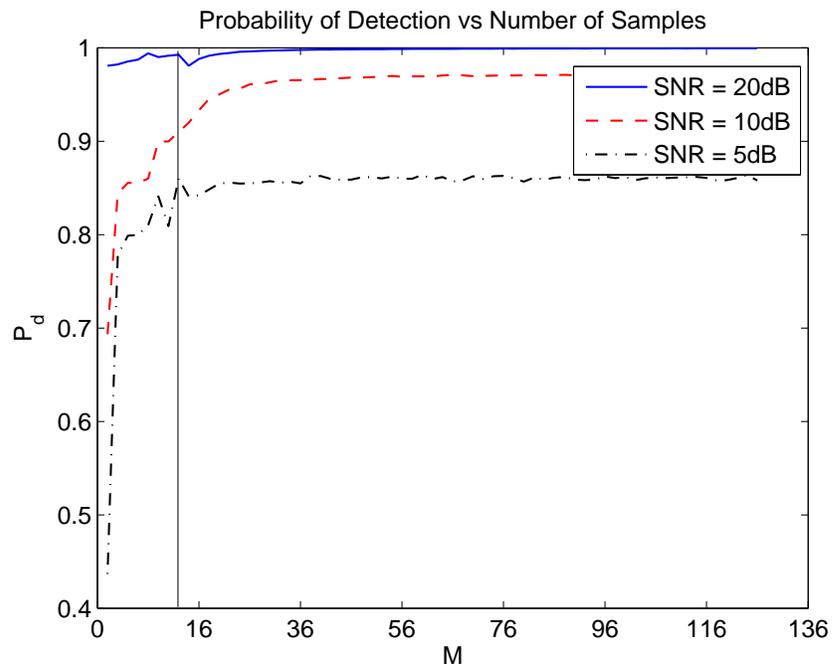


Figure 6.3: Probability of Detection versus Number of Samples.

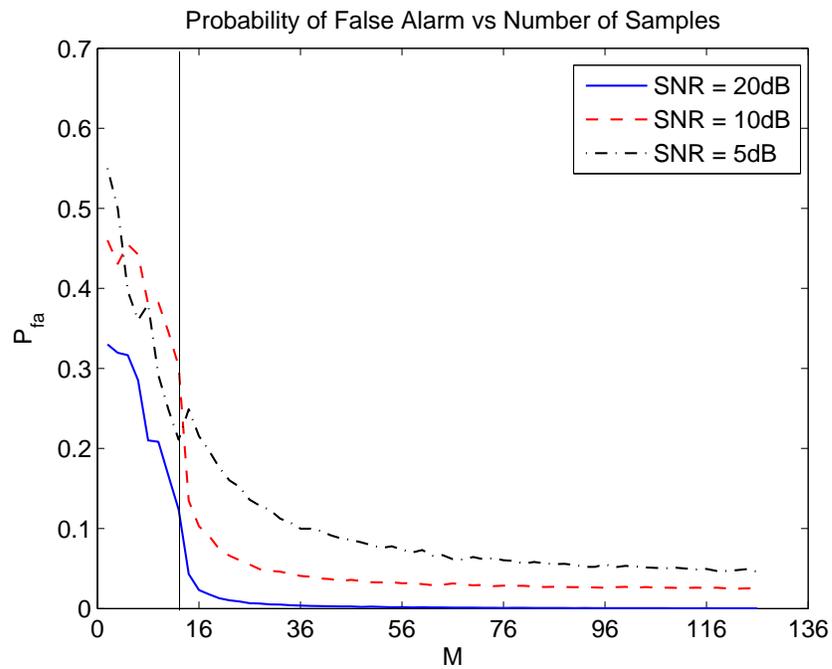


Figure 6.4: Probability of False Alarm versus Number of Samples.

in these figures corresponds to the boundary between the sample rich (right side of boundary) and sample poor (left side of boundary) cases. It is interesting to note that at the boundary point where sample support is poor but the number of eigenvalues is still 12, the eigenvalue gap and J-divergence reach their maximum values. Thus, as sample support decreases the J-divergence becomes a superficial measure for separability of the two hypotheses. Figures 6.3 and 6.4 illustrate the fact that the number of misdetections and false alarms also increase as the sample support decreases. The decline in the J-divergence in sample poor region is due to the fact that the J-divergence is computed with lesser number of eigenvalues since the rank of  $\hat{S}$  decreases as the number of samples decreases. It can also be noted that the eigenvalue gap and the J-divergence reached a steady state value (theoretical value for the specific cases considered) after the number of samples passes 40.

Comparing the three different SNR scenarios, several interesting observations can be made. First, the eigenvalue gap is lower for a higher SNR at all sample support sizes. Additionally, they all reach their steady state values with the lowest for SNR=20dB. To explain this behavior, we note that the eigenvalues are computed from the “signal-to-noise ratio” matrix,  $\hat{S}$  and if  $\gamma_1$  and  $\gamma_{12}$  are the largest and smallest eigenvalues of the signal covariance matrix  $R_s$ , it can easily be shown that the eigenvalue gap is  $(\lambda_1/\lambda_{12}) = (\gamma_1 + \sigma_n^2)/(\gamma_{12} + \sigma_n^2) \approx (SNR+1)/(\gamma_{12}/\sigma_n^2 + 1)$ , where it is assumed that the covariance matrix of additive noise is  $R_0 = \sigma_n^2 I$ . Since the eigenvalues of  $R_s$  are fixed as the  $(SNR \approx \gamma_1/\sigma_n^2)$  decreases the eigenvalue gap will increase depending on the relative values of  $\sigma_n^2$  and  $\gamma_{12}$ . Second, the J-divergence reaches steady-state values for all three SNR cases, when the number of samples  $\geq 40$ . However, the steady-state value of J-divergence is much larger for higher SNR’s. This results is consistent with those in Figures 6.3 and 6.4 which show dramatic increase in the averaged  $P_d$  and decrease in the averaged  $P_{fa}$  when SNR was higher. The deterioration in the overall detection performance when SNR decreases is more prominent

when the sample support  $\leq 40$  i.e. in the region where J-divergence is flawed. This implies that as SNR decreases more samples is required to distinguish between the two hypotheses.

## 6.5 Conclusion

In this Chapter, we developed the sampled version of the standard Gauss-Gauss detector and studied the effect of sample support size on the eigenvalues of the signal-to-noise ratio matrix. More specifically, two different cases were investigated namely, the sample rich case, in which the number of samples is greater or equal to the dimension of the channel, and the sample poor case, in which the number of samples is less than the dimension of the channel. The sampled version of the linear detector was then extended to the nonlinear kernel case where the data vectors are mapped into a higher dimensional feature space. The centered version of the nonlinear kernel detector was also derived. The nonlinear kernel Gauss-Gauss detector can become intractable due to high or *possibly infinite* dimension of the feature space comparing to the dimension of the original signal space.

It was noted that as the sample support decreases the J-divergence measure becomes defective as a global measure of separability between the two hypotheses. The simulation results for sample support sizes ranging from 4 to 256 samples and different SNR cases indicated that as the number of samples decreases toward the dimension of the data vector the J-divergence superficially increases, whereas the overall  $P_d$  declined. When sample size becomes  $\geq 40$  the eigenvalue gap and J-divergence reached their steady state values for all the three SNR cases considered. For lower SNR, J-divergence reached a lower steady-state value. While, the averaged  $P_d$  decreased and the averaged  $P_{fa}$  increased, especially when sample support is  $\leq 40$ . This implies that as SNR decreases, more samples are required to distinguish between the two hypotheses. Therefore, in order to use the kernel Gauss-Gauss detector for any case

an adequate number of samples is required for implementation.

# CHAPTER 7

## CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

### 7.1 Conclusions and Discussions

A new coherence-based method for the detection of underwater objects from sonar imagery was developed in this thesis in order to provide a mechanism for robust target detection in changing environmental conditions. Using the developed coherence-based detector three detection schemes were developed to address three specific problems, namely, the single sonar, the dual disparate sonar, and the distributed sonar platform detection problems. The single sonar detection system utilizes coherence properties in one sonar image to detect a target ROI. The dual disparate sonar detection system utilizes two sonars operating at disparate frequencies and resolutions and provides coherence-based detection based upon two ROI's within the co-registered images. Finally, the distributed detection system utilizes several dual disparate detection systems as local decision makers, and a fusion center to generate a final detection decision based upon three or more co-registered sonar images. The use of multiple local decision makers allows generating multiple decisions about the environment and by sending only the local decisions to the fusion center the amount of information that needs to be transmitted in a NSA setup is substantially reduced. Each of these detection systems exploits the coherence (or common) information in one image or between the co-registered images using the two-channel CCA to make a high confidence decision about the presence of an object. These systems were then compared, which led to the distributed detection system providing the best overall detection

performance.

Considering that each of the detection systems needs to extract the coherent information from one or multiple sonar images, CCA is considered to be the optimal tool for coherence analysis. Because CCA transforms the two-channel data into canonical coordinates that have maximal correlation between individual pairs of coordinates in comparison to any other linear mapping. This method provides a perfect framework for relating the columns in an ROI for the single sensor case and between two co-registered ROI's in the dual disparate sonar case. Moreover, once the common features between the two sonar images are identified, the *coherence pattern* represented by the canonical correlations can be used for subsequent classification.

The performance of each of these coherence-based detection schemes was tested on the Sonar8 and multi-platform sonar data sets, each of which contains targets in background with varying density of clutter. The Sonar8 data set contains high resolution side-scan sonar imagery from one sensor. The images contained in this data set consisted of envelope data (magnitude of the complex image) that is the output of the beamformer. The multi-platform sonar data set contains a high frequency high resolution side-looking sonar image over the target field and three broadband sonar images co-registered over the same region on the seafloor. The images in this data set are complex and are the direct output of the beamformer.

Data preprocessing and coherence-based feature extraction using CCA were also discussed. Data preprocessing was necessary in order to help reduce clutter and enhance the signature of the targets in a sonar image and to prepare the sonar imagery for detection. In order for CCA to be successfully applied to the data contained in the ROI, the ROI needed to be *channelled*, *vectored*, and *averaged* to optimally extract the coherent information. Different channeling methods were used for the single sensor case and the dual disparate sensor case due to the different application of the coherence-based detector in each case. Specifically, coherence was extracted between

consecutive columns in an ROI for the single sonar application where the  $\mathbf{x}$ - and  $\mathbf{y}$ -channels were chosen to be consecutive columns in the ROI. For the dual disparate case, coherence was extracted between a pair of blocks in a pair of co-registered ROI's and a block from one ROI formed the  $\mathbf{x}$ -channel while the corresponding block from the other ROI formed the  $\mathbf{y}$ -channel.

The standard Gauss-Gauss detector that was cast in the CCA framework [29] cannot be directly applied to the problem presented in this thesis, due to the fact that in the proposed applications the hypotheses involve a composite channel. This changes the covariance matrices under the two hypotheses and changes the formulation of the standard Gauss-Gauss detector. In this thesis, we presented a new formulation for the log-likelihood and J-divergence for the new hypotheses making the detector suitable for detection of objects from sonar imagery. Specifically we demonstrated a new way in finding the inverses and relating the composite two-channel hypothesis problem back to the standard problem. Through this development we showed the relationship to the standard CCA-based detector in [29] for ease in implementation.

Using this formulation, the detection scheme was then applied to both single sonar and the dual disparate sonar cases. The single sonar case utilizes one high resolution high frequency sonar image. Using the extracted canonical coordinates and corresponding correlations the log-likelihood was computed for every pair of columns in an ROI and if more than 50% of the columns have log-likelihood values above a preselected detection threshold a detection is declared for that ROI. This system was tested on the Sonar8 data set. The detection results on this data set demonstrated that the system detected 275 out of the 286 targets with an average of 24 false alarms per image. Moreover, at the knee point of the ROC curve for the easy cases (less clutter density) we had  $P_d = 89\%$  and  $P_{fa} = 11\%$ . For the medium cases (medium clutter density), the knee point was  $P_d = 84\%$  and  $P_{fa} = 16\%$  and for hard cases (very high clutter density), the knee point of the ROC gave  $P_d = 56\%$  and  $P_{fa} = 44\%$ .

The dual disparate sensor case utilizes the advantages of using side-scan sonar imagery with disparate frequency and resolution characteristics. The use of multiple sonar types allows one to use a high resolution sonar image with good target definition and a low resolution broadband sonar image with good clutter suppression abilities co-registered over the same region on the seafloor. Using the extracted canonical coordinates and corresponding correlations the log-likelihood was computed for every pair of blocks within the ROI's of the co-registered sonar images and if more than 50% of the blocks have log-likelihood values that fall above the preselected threshold a detection was declared for that pair of ROI's. This system was tested on the multi-platform sonar data set provided by the NSWC-PC. Using a high frequency and a broadband sonar images the detector detected 51 of the 53 targets and had an average of 10 false alarms per image. The knee point of the ROC gave  $P_d = 95\%$  and  $P_{fa} = 5\%$ . The joint information from the broadband sonar and high frequency sonar led to overall detection results that were very good.

Lastly, the dual disparate detector was used in the distributed detection system where more than one dual disparate sonar platform is used. The decisions of the individual dual disparate detectors are then fused to yield a high confidence final decision. The fusion rule is based on the work in [48] and is modified to utilize an observation from the environment in the fusion center. This observation helps to greatly increase the probability of detection and decrease the false alarm rate. The implemented system consisted of one dual disparate detector utilizing two broadband images and a second dual disparate detector utilizing one of the same broadband images and a high frequency sonar image. In the first detector, the coherence information in the two broadband sonar images is exploited to suppress clutter ROI's and detect potential targets. The coherence information in the second dual disparate detector solidifies the decision in the first detector by verifying the joint presence in the high-frequency and broadband images. The decisions from the two detectors are sent to the fusion

center where a final decision is made using the preliminary decisions and the fusion center's own observation. When the same multi-platform sonar data set was applied to the distributed detection system that used three sonar images it was shown that the fusion process significantly improved the probability of correct detection while reducing the incidence of a false alarm. At the knee point of the ROC, the system provided a  $P_d = 99\%$  and  $P_{fa} = 1\%$ , and all of the targets were detected successfully (hard-limiting threshold) with an average of only 7 false alarms per image on the entire NSWC multi-platform sonar database.

The last study conducted in this thesis was on the effect of sample support on the Gauss-Gauss detector and its extension to the kernel (non-linear) case. More specifically, we developed the sampled version of the standard Gauss-Gauss detector and studied the effects of sample support size on the eigenvalues of the signal-to-noise ratio matrix. Two different cases were investigated namely, the sample rich case, in which the number of samples is greater or equal to the dimension of the data channel, and the sample poor case, in which the number of samples is less than the dimension of the data channel. From the sampled version, the kernel version of the Gauss-Gauss detector was derived as well as the centered version of the nonlinear kernel detector. It was noted that as the sample support decreases the J-divergence measure becomes defective as a global measure of separability between the two hypotheses. The simulation results indicated that as the number of samples decreases toward the dimension of the data vector the J-divergence superficially increases, whereas the overall  $P_d$  declined. It was also found that for lower SNR, J-divergence reached a lower steady-state value while, the averaged  $P_d$  decreased and the averaged  $P_{fa}$  increased. This implied that as SNR decreases, more samples are required to distinguish between the two hypotheses.

## 7.2 Future Work

Although, the coherence-based detectors proposed in this thesis offer powerful tools for detection of underwater targets from multiple disparate sonar platforms, there are several important areas and extensions that can be pursued in the future. These include, but are not limited to:

- Extension of the two-platform Gauss-Gauss detector cast in the CCA framework to the multi-platform case using the multi-channel coherence analysis (MCA) [59] and developing the corresponding Gauss-Gauss detector in the MCA framework. The MCA method exploits the coherence between 2 or more channels or sonar images lending itself perfectly to detection using multiple disparate platforms. Preliminary studies presented in [60] attest to the fact that MCA-based detection can provide the same detection performance as that of the distributed collaborative detector in Chapter 5 with much lesser structural complexity. Future research is needed to thoroughly study the performance of the detector on more multi-platform sonar data sets.
- Extension of the coherence-based detection method to account for the multi-hypothesis testing in the disparate sensor detection problem. For the two-platform cases two additional hypotheses must be added in the formulations to detect the targets that don't appear very well in either the high-frequency or broadband sonar images. This requires extending the Gauss-Gauss detector to M-hypothesis testing [61] problem and relating the corresponding J-divergence detectability measure to the canonical correlations of the two sonar imagery data.
- The distributed detection implementation could be extended to include collaboration between the local decision makers similar to that of the collaborative

multi-aspect classifier in [33]. The next step would be to develop a new collaborative distributed detection methodology that takes into account (a) limited communication bandwidth for communicating essential target information among multiple disparate AUV's and the mother ship; (b) computational limitations of the processing boards on each AUV platform; (c) near real-time decision-making requirement; and (d) practically feasible, versatile and robust implementation of the coherence-based detection and feature extraction methods developed in this thesis.

- The data used in this study was limited to only a few runs and types of underwater targets. Ideally, the next step in the development of the coherence-based detector would be to test the performance on more data to prove the usefulness of the detection systems developed in this thesis. The testing on more difficult data sets provided by the NSWC as well those including more man-made non-targets will be done in the future. More specifically, a study on the effect of different bottom types, target orientations, sonar aspect, resolution, and SNR on the probability of detection and false alarm rate would be insightful and help to illustrate the real effectiveness of the detector for realistic underwater target detection problems.
- The main development of this thesis was on the detection of underwater targets from sonar imagery and not much work was carried out on the classification. Another potential extension of this research would be to study the use of canonical correlation features for classification of targets and non targets. If successful, this will allow us to carry out simultaneous detection and classification using only the extracted canonical coordinates and correlations without requiring a separate feature extraction system.
- The coherence-based detector developed in this thesis is applicable not only to

sonar image detection, but could be used on other disparate sensory systems, i.e. magnetic, infrared, and optical. A study of its usefulness on these types of sensing modalities would highly be valuable. By finding the coherence information between more than one type of sensors the detection and classification performance could be improved.

- Although, the coherence-based detector developed in this thesis lead to a small number of false alarms comparing to other detectors, we believe the Gaussianity assumption inherent in the standard CCA method does not allow for better representation of the background clutter in the design of the detector. This is due to the fact that studies [62] - [64] have indicated that different bottom types exhibit different distributions depending on the grazing angle, sonar frequency and range and these distributions are typically non-Gaussian. An extension of the coherence analysis framework to account for the non-Gaussian distribution of the background clutter would need to be developed. More specifically, new methods would need to be developed that exploit higher order statistical properties of the background clutter. From these methods new formulations for the log-likelihood and J-divergence for the non-Gaussian case would need to be developed to extend the developed multi-platform detection algorithm to account for the non-Gaussian background clutter.
- Another possible extension to the target detection from sonar imagery is multiple hypothesis testing. There has been a large amount of research devoted to multiple testing, especially in the areas of bioinformatics, genomics, and brain imaging [65], [66]. The idea behind a multiple testing approach is that by performing multiple statistical tests where a number of hypotheses is rejected and another number of hypotheses is accepted. By making a small number of false discoveries and by controlling a suitable error rate one can maximize the power

of each test at the same time and thus reduce the overall false alarm rate.

## REFERENCES

- [1] G. J. Dobeck, “Image normalization using the serpentine forward-backward filter: application to high-resolution sonar imagery and its impact on mine detection and classification,” *Proc. SPIE*, vol. 5734, pp. 90–110, April 2005.
- [2] M. R. Azimi-Sadjadi and J. D. Tucker, “Target detection from dual disparate sonar platforms using canonical correlations,” *Proc. SPIE*, vol. 6953, pp. J1 – J10, March 2008.
- [3] T. Aridgides, P. Libera, M. Fernandez, and G. J. Dobeck, “Adaptive filter/feature orthogonalization processing string for optimal LLRT mine classification in side-scan sonar imagery,” *Proc. SPIE*, vol. 2765, pp. 110–121, April 1996.
- [4] G. J. Dobeck, J. Hyland, and L. Smedley, “Automated detection/classification of sea mines in sonar imagery,” *Proc. SPIE*, vol. 3079, pp. 90–110, April 1997.
- [5] G. J. Dobeck, “Fusing sonar images for mine detection and classification,” *Proc. SPIE*, vol. 3710, pp. 602–614, April 1999.
- [6] C. Ciany and J. Huang, “Data fusion of VSW CAD/CAC algorithms,” *Proc. SPIE*, vol. 4038, pp. 413–420, April 2000.
- [7] C. M. Ciany and J. Huang, “Computer aided detection/computer aided classification and data fusion algorithms for automated detection and classification of underwater mines,” *Proc. of MTS/IEEE OCEANS 2000 Conference*, vol. 1, pp. 277–284, 2000.

- [8] C. M. Ciany, W. C. Zurawski, G. J. Dobeck, and D. R. Weilert, "Real-time performance of fusion algorithms for computer aided detection and classification of bottom mines in the littoral environment," *Proc. of MTS/IEEE OCEANS 2003 Conference*, vol. 2, pp. 1119–1125, Sept. 2003.
- [9] V. Chandran, S. Elgar, and A. Nguyen, "Detection of mines in acoustic images using higher order spectral features," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 3, pp. 610–618, July 2002.
- [10] T. Aridgides, M. F. Fernandez, and G. J. Dobeck, "Adaptive three-dimensional range-cross-range-frequency filter processing string for sea mine classification in side-scan sonar imagery," *Proc. SPIE*, vol. 3079, pp. 111–122, April 1997.
- [11] T. Aridgides and M. F. Fernandez, "Automated target classification in high resolution dual frequency sonar imagery," *Proc. SPIE*, vol. 6553, pp. 1–12, April 2007.
- [12] T. Aridgides and M. F. Fernandez, "Enhanced ATR algorithm for high resolution multi-band sonar imagery," *Proc. SPIE*, vol. 6953, pp. 0–10, April 2008.
- [13] T. Aridgides and M. F. Fernandez, "Automatic target recognition algorithm for high resolution multi-band sonar imagery," *Proc. of MTS/IEEE Oceans 2008 Conference*, to be published.
- [14] J. D. Tucker, M. R. Azimi-Sadjadi, and G. J. Dobeck, "Coherent-based method for detection of underwater objects from sonar imagery," *Proc. SPIE*, vol. 6553, pp. U1–U8, April 2007.
- [15] J. D. Tucker, M. R. Azimi-Sadjadi, and G. J. Dobeck, "Canonical coordinates for detection and classification of underwater objects from sonar imagery," *Proc. of IEEE OCEANS 2007 Conference Europe*, pp. 1–6, June 2007.

- [16] J. D. Tucker, N. Klausner, and M. R. Azimi-Sadjadi, "Target detection in m-disparate sonar platforms using multichannel hypothesis testing," *Proc. of MTS/IEEE Oceans 2008 Conference*, to be published.
- [17] C. Ciany and W. Zurawski, "Impact of image normalization and quantization on the performance of sonar computer-aided detection/computer-aided classification (CAD/CAC) algorithms," *Proc SPIE*, vol. 6553, pp. 1T–9T, 2007.
- [18] R. O. Duda, P. E. Hart, and D. G. Sotrk, *Pattern Classification*. John Wiley and Sons, 2001.
- [19] T. Aridgides and M. Fernandez, "Enhanced ATR algorithm for high resolution multi-band sonar imagery," *Proc. SPIE*, vol. 6953, pp. 0–1, March 2008.
- [20] L. Atallah, C. Shang, and R. Bates, "Object detection at different resolution in archaeological side-scan sonar images," *Proc. of IEEE Oceans 2005 Conference Europe*, vol. 1, pp. 287–292, June 2005.
- [21] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July,October 1948.
- [22] T. Kadir and M. Brady, "Saliency, scale and image description," *International Journal of Computer Vision*, vol. 45, no. 2, pp. 83–105, November 2001.
- [23] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 1, pp. 90–105, Jan 2003.
- [24] R. Kindermann and J. L. Snell, *Markov Random Fields and their Applications*. AMS, 1980.

- [25] C. Chesnaud, P. Refregier, , and V. Boulet, “Statistical region snake-based segmentation adapted to different physical noise models,” *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 1145–1157, 1999.
- [26] E. Dura, Y. Zhang, X. Liao, G. J. Dobeck, and L. Carin, “Active learning for detection of mine-like objects in side-scan sonar imagery,” *IEEE Journal of Oceanic Engineering*, vol. 30, no. 2, pp. 360–371, April 2005.
- [27] L. Scharf and C. Mullis, “Canonical coordinates and the geometry of inference, rate, and capacity,” *IEEE Transactions on Signal Processing*, vol. 48, no. 3, pp. 824–891, March 2000.
- [28] H. Hotelling, “Relations between two sets of variates,” *Biometrika*, vol. 28, pp. 321–377, 1936.
- [29] A. Pezeshki, L. L. Scharf, J. K. Thomas, and B. D. Van Veen, “Canonical coordinates are the right coordinates for low-rank Gauss-Gauss detection and estimation,” *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4817–4820, Dec 2006.
- [30] A. Pezeshki, M. Azimi-Sadjadi, L. Scharf, and M. Robinson, “Underwater target classification using canonical correlations,” *Proc. of MTS/IEEE Oceans 2003 Conference*, vol. 4, pp. 1906–1911, Sept 2003.
- [31] A. Pezeshki, M. R. Azimi-Sadjadi, L. L. Scharf, and M. Robinson, “Underwater target classification using canonical correlations,” *Proc. of MTS/IEEE Oceans 2003 Conference*, pp. 1906–1911, September 2003.
- [32] A. Pezeshki, M. R. Azimi-Sadjadi, and L. L. Scharf, “Undersea target classification using canonical correlation analysis,” *IEEE Journal of Oceanic Engineering*, vol. 32, no. 4, pp. 948–955, Oct. 2007.

- [33] J. Cartmill, N. Wachowski, and M. R. Azimi-Sadjadi, “Buried underwater object classification using a collaborative multi-aspect classifier,” *IEEE Journal of Oceanic Engr.*, vol. 34, no. 1, pp. 32–44, 2009.
- [34] N. Wachowski and M. R. Azimi-Sadjadi, “Buried underwater target classification using frequency subband coherence analysis,” *Proc. of MTS/IEEE Oceans 2008 Conference*, to be published.
- [35] L. L. Scharf and B. D. Van Veen, “Low rank detectors for Gaussian random vectors,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, no. 11, pp. 1579–1582, Nov 1987.
- [36] B. Schölkopf and A. J. Somola, *Learning with Kernels*. The MIT press, 2002.
- [37] A. Ruiz and P. E. Lopez-de-Teruel, “Nonlinear kernel-based statistical pattern analysis,” *IEEE Trans. on Neural Networks*, vol. 12, pp. 16–31, 2001.
- [38] B. Schölkopf, C. J. C. Burges, and A. J. Somola, *Advances in Kernel Methods*. The MIT press, 1999.
- [39] N. A. Rimski-Korsakov, Y. S. Russak, and R. B. Pavlov, “Simple digital system for side-scan sonar data imaging,” *Proc. of MTS/IEEE Oceans 1994 Conference*, vol. 1, pp. 643–645, Sept 1994.
- [40] W. H. Key, “Side scan sonar technology,” *Proc. of MTS/IEEE Oceans 2000 Conference*, vol. 2, pp. 1029–1033, Sept 2000.
- [41] R. Bamler, “A comparison of range-Doppler and wavenumber domain SAR focusing algorithms,” *IEEE Trans. on Geoscience and Remote Sensing*, vol. 30, no. 4, pp. 706–713, 1992.
- [42] M. Soumekh, *Fourier Array Imaging*. Prentice Hall, Englewood Cliffs, NJ, 1994.

- [43] H. L. Van Trees, *Optimum Array Processing*. Wiley-Interscience, 2002.
- [44] R. H. Stolt, "Migration by Fourier transform," *Geophysics*, vol. 43, no. 1, pp. 23–48, 1978.
- [45] S. C. Butler, "Triply resonant broadband transducers," *Proc. of MTS/IEEE Oceans 2002 Conference*, vol. 4, pp. 2334–2341, 29-31 Oct. 2002.
- [46] S. C. Butler, "High frequency triply resonant broadband transducer array development at NUWC," *Proc. of MTS/IEEE OCEANS 2003 Conference*, vol. 5, pp. 2380–5, 22-26 Sept. 2003.
- [47] H. L. Van Trees, *Detection, Estimation, and Modulation Theory Part I*. John Wiley and Sons, 1968.
- [48] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer-Verlag New York, 1996.
- [49] M. L. Eaton, *Multivariate Statistics: A Vector Space Approach*. New York: Wiley, 1983, ch. 10.
- [50] E. L. Lehman, *Testing Statistical Hypotheses*. New York: Wiley, 1986.
- [51] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third ed. John Hopkins University Press, 1996.
- [52] S. Haykin, *Neural Networks*, 2nd ed. Prentice-Hall, Inc. Upper Saddle River, NJ, 1999.
- [53] D. F. Specht, "Probabilistic neural network," *Neural Networks*, vol. 3, pp. 109–118, 1990.

- [54] D. F. Specht and P. D. Shapiro, "Generalization ability of probabilistic neural networks compared with backpropagation networks," *Proceedings of the 1991 IEEE International Joint Conference on Neural Networks*, pp. 458–461, 1991.
- [55] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math Statist.*, vol. 33, pp. 1065–1076, 1962.
- [56] N. M. Nasrabadi and H. Kwon, "Kernel spectral matched filter for hyperspectral target detection," *Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Processing (ICASSP '05)*, vol. 4, pp. 18–23, March 2005.
- [57] H. Kwon and N. M. Nasrabadi, "Kernel RX-algorithm: A nonlinear anomaly detector for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 43, no. 2, pp. 388–397, Feb 2005.
- [58] H. Kwon and N. M. Nasrabadi, "Kernel matched subspace detectors for hyperspectral target detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 2, pp. 178–194, Feb. 2006.
- [59] B. D. Thompson and M. R. Azimi-Sadjadi, "Iterative multi-channel coherence analysis with applications," *Neural Networks*, vol. 21, no. 2-3, pp. 493–501, 2008.
- [60] N. Klausner, M. R. Azimi-Sadjadi, and J. D. Tucker, "Underwater target detection from multi-platform sonar imagery using multi-channel coherence analysis," *Proc. of SMC 2009 Conference*, to be published.
- [61] E. Drakopoulos, J. J. Chao, and C. C. Lee, "A two-level distributed multiple hypothesis decision system," *IEEE Trans. on Automatic Control*, vol. 37, no. 3, pp. 380–384, Mar 1992.
- [62] S. Stanic and E. Kennedy, "Reverberation fluctuation from a smooth seafloor," *IEEE Journal of Oceanic Engr.*, vol. 18, pp. 95–99, 1993.

- [63] S. Stanic, R. Goodman, K. Briggs, N. P. Chotiros, and E. Kennedy, "Shallow-water bottom reverberation measurements," *IEEE Journal of Oceanic Engr.*, vol. 23, pp. 203–210, 1998.
- [64] J. T. Cobb and K. C. Slatton, "A parameterized statistical sonar image texture model," *Proc. SPIE*, vol. 6953, pp. K1–K12, 2008.
- [65] S. Zhong, L. Tian, C. Li, K. F. Storch, and W. H. Wong, "Comparative analysis of gene sets in the gene ontology space under the multiple hypothesis testing framework," *Proc. of IEEE Computational Systems Bioinformatics Conf.*, pp. 425–435, Aug. 2004.
- [66] A. Farcomeni, "A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion," *Statistical Methods in Medical Research*, vol. 17, no. 4, pp. 347–388, 2008.
- [67] A. Pezeshki, L. L. Scharf, M. R. Azimi-Sadjadi, and Y. Hua, "Two-channel constrained least squares problems: Solutions using power methods and connections with canonical coordinates," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 121–135, Jan 2005.

# APPENDIX A

## CANONICAL CORRELATION ANALYSIS REVIEW

In this appendix, we provide a review of the CCA method in which a set of basis vectors is found for two sets of multidimensional variables such that correlations between the projections onto these basis vectors are mutually maximized. The material presented here, as well as much of the language and terminology, are drawn from [27]. CCA was proposed by Hotelling [28] for the analysis of linear dependence between two data channels. CCA decomposes the linear dependence between the original channels into the linear dependence between the canonical coordinates of the channels, where this linear dependence is easily determined by the corresponding canonical correlations [27].

Consider the composite data vector  $\mathbf{z}$  consisting of two random vectors  $\mathbf{x} \in \mathbb{R}^m$  and  $\mathbf{y} \in \mathbb{R}^n$ , i.e.

$$\mathbf{z} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \in \mathbb{R}^{(m+n)}. \quad (\text{A-1})$$

For the remainder of the derivations, it is assumed that  $m \geq n$ , also the notation  $(\cdot)^H$  represents the Hermitian operation. Assume that  $\mathbf{x}$  and  $\mathbf{y}$  have zero means and share the composite covariance matrix

$$R_{zz} = E[\mathbf{z} \mathbf{z}^H] = E \left[ \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \begin{pmatrix} \mathbf{x}^H & \mathbf{y}^H \end{pmatrix} \right] = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix}. \quad (\text{A-2})$$

If  $\mathbf{x}$  and  $\mathbf{y}$  are now replaced by their corresponding whitened vectors, then the

composite vector  $\boldsymbol{\xi}$  is

$$\boldsymbol{\xi} = \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\nu} \end{bmatrix} = \begin{bmatrix} R_{xx}^{-1/2} & \mathbf{0} \\ \mathbf{0} & R_{yy}^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \quad (\text{A-3})$$

where  $R_{xx}^{1/2}$  is a square-root of  $R_{xx}$  with  $R_{xx}^{1/2} R_{xx}^{H/2} = R_{xx}$  and  $R_{xx}^{-1/2} R_{xx} R_{xx}^{-H/2} = \mathbf{I}$ .

The covariance matrix of  $\boldsymbol{\xi}$  may be written as

$$R_{\xi\xi} = E[\boldsymbol{\xi} \boldsymbol{\xi}^H] = E \left[ \begin{pmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\nu} \end{pmatrix} \begin{pmatrix} \boldsymbol{\zeta}^T & \boldsymbol{\nu}^H \end{pmatrix} \right] = \begin{bmatrix} R_{\zeta\zeta} & R_{\zeta\nu} \\ R_{\nu\zeta} & R_{\nu\nu} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{C} \\ \mathbf{C}^H & \mathbf{I} \end{bmatrix}, \quad (\text{A-4})$$

where

$$\mathbf{C} = E[\boldsymbol{\zeta} \boldsymbol{\nu}^T] = E[(R_{xx}^{-1/2} \mathbf{x})(R_{yy}^{-1/2} \mathbf{y})^H] = R_{xx}^{-1/2} R_{xy} R_{yy}^{-H/2} \quad (\text{A-5})$$

is called the *coherence matrix* of  $\mathbf{x}$  and  $\mathbf{y}$  [27], [67]. Therefore, the coherence matrix  $\mathbf{C}$  is the cross-covariance matrix between the whitened versions of  $\mathbf{x}$  and  $\mathbf{y}$ . Correspondingly, the coordinates  $\boldsymbol{\zeta}$  and  $\boldsymbol{\nu}$  are called the *coherence coordinates*. Now it is possible to determine the singular value decomposition (SVD) of the coherence matrix, namely

$$\begin{aligned} \mathbf{C} &= R_{xx}^{-1/2} R_{xy} R_{yy}^{-H/2} = \mathbf{F} \mathbf{K} \mathbf{G}^H \quad \text{and} \\ \mathbf{F}^H \mathbf{C} \mathbf{G} &= \mathbf{F}^H R_{xx}^{-1/2} R_{xy} R_{yy}^{-T/2} \mathbf{G} = \mathbf{K}, \end{aligned} \quad (\text{A-6})$$

where  $\mathbf{F} \in \mathbb{R}^{m \times m}$  and  $\mathbf{G} \in \mathbb{R}^{n \times n}$  are orthogonal matrices [51], i.e.

$$\mathbf{F}^H \mathbf{F} = \mathbf{F} \mathbf{F}^H = \mathbf{I}(m) \quad \text{and} \quad \mathbf{G}^H \mathbf{G} = \mathbf{G} \mathbf{G}^H = \mathbf{I}(n), \quad (\text{A-7})$$

and

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}(n) \\ \mathbf{0} \end{bmatrix} \in \mathbb{R}^{m \times n} \quad (\text{A-8})$$

is a diagonal singular value matrix, with  $\mathbf{K}(n) = \text{diag}[k_1, k_2, \dots, k_n]$  and  $1 \geq k_1 \geq k_2 \geq \dots \geq k_n > 0$ .

We then use the orthogonal matrices  $F$  and  $G$  to transform the whitened composite vector  $\boldsymbol{\xi}$  into the canonical composite vector  $\mathbf{w}$ ,

$$\mathbf{w} = \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} F^H & \mathbf{0} \\ \mathbf{0} & G^H \end{bmatrix} \begin{bmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\nu} \end{bmatrix} = \begin{bmatrix} F^H & \mathbf{0} \\ \mathbf{0} & G^H \end{bmatrix} \begin{bmatrix} R_{xx}^{-1/2} & \mathbf{0} \\ \mathbf{0} & R_{yy}^{-1/2} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}. \quad (\text{A-9})$$

Then, the covariance matrix for the canonical composite vector  $\mathbf{w}$  is obtained as

$$R_{ww} = E[\mathbf{w}\mathbf{w}^H] = E \left[ \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix} (\mathbf{u}^H \quad \mathbf{v}^H) \right] = \begin{bmatrix} R_{uu} & R_{uv} \\ R_{vu} & R_{vv} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{K}^H & \mathbf{I} \end{bmatrix}. \quad (\text{A-10})$$

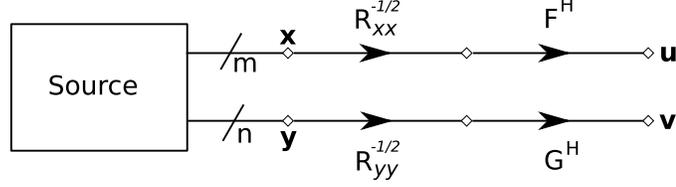
The elements of  $\mathbf{u} = [u_i]_{i=1}^m \in \mathbb{R}^m$  are referred to as the *canonical coordinates* of  $\mathbf{x}$  and the elements of  $\mathbf{v} = [v_i]_{i=1}^n \in \mathbb{R}^n$  are the canonical coordinates of  $\mathbf{y}$ . The diagonal cross-correlation matrix  $\mathbf{K}$ ,

$$\mathbf{K} = E[\mathbf{u}\mathbf{v}^H] = E[(F^H R_{xx}^{-1/2} \mathbf{x})(G^H R_{yy}^{-1/2} \mathbf{y})^H] = F^H \mathbf{C} G \quad (\text{A-11})$$

is called the *canonical correlation matrix* of *canonical correlations*  $k_i$ , with  $1 \geq k_1 \geq k_2 \geq \dots \geq k_n > 0$ . Thus, the canonical correlations measure the correlations between pairs of corresponding canonical coordinates. That is,  $E[u_i v_j] = k_i \delta_{ij}$ ;  $i \in [1, n]$ ,  $j \in [1, m]$ , with  $\delta_{ij}$  being the Kronecker delta. The canonical correlations  $k_i$  are also the singular values of the coherence matrix  $\mathbf{C}$ . Correspondingly,  $\mathbf{K}\mathbf{K}^H$  is the squared canonical correlation matrix of the squared canonical correlations  $k_i^2$ . Since  $F$  and  $G$  are orthogonal matrices, we may write the squared coherence matrix  $\mathbf{C}\mathbf{C}^H$  as

$$\begin{aligned} \mathbf{C}\mathbf{C}^H &= R_{xx}^{-1/2} R_{xy} R_{yy}^{-1} R_{yx} R_{xx}^{-H/2} \\ &= \mathbf{F} \mathbf{K} \mathbf{G}^H \mathbf{G} \mathbf{K}^H \mathbf{F}^H = \mathbf{F} \mathbf{K} \mathbf{K}^H \mathbf{F}^H. \end{aligned} \quad (\text{A-12})$$

This shows that the squared canonical correlations  $k_i^2$  are the eigenvalues of the squared coherence matrix  $\mathbf{C}\mathbf{C}^H$ , or equivalently, of the matrix  $R_{xx}^{-H/2} \mathbf{C}\mathbf{C}^H R_{xx}^{H/2} = R_{xx}^{-1} R_{xy} R_{yy}^{-1} R_{yx}$ . It is interesting to note that these eigenvalues are invariant to the choice of a square-root for  $R_{xx}$ .



**Figure A-1:** Transformation from standard coordinates  $\mathbf{x}$  and  $\mathbf{y}$  to canonical coordinates  $\mathbf{u}$  and  $\mathbf{v}$ .

Figure A-1 illustrates the transformation from standard coordinates  $\mathbf{x}$  and  $\mathbf{y}$  to coherence coordinates  $\boldsymbol{\zeta}$  and  $\boldsymbol{\nu}$  and then to canonical coordinates  $\mathbf{u}$  and  $\mathbf{v}$ . It can be noted that the transformation from standard coordinates  $\mathbf{x}$  and  $\mathbf{y}$  to canonical coordinates  $\mathbf{u}$  and  $\mathbf{v}$  can be represented by  $\mathbf{u} = \mathbf{W}^H \mathbf{x}$  and  $\mathbf{v} = \mathbf{D}^H \mathbf{y}$  where  $\mathbf{W}^H = \mathbf{F}^H \mathbf{R}_{xx}^{-1/2}$  and  $\mathbf{D}^H = \mathbf{G}^H \mathbf{R}_{yy}^{-1/2}$ . In this case,  $\mathbf{W}$  and  $\mathbf{D}$  are known as the canonical mapping matrices.

The canonical correlations  $k_i$  are invariant to block-diagonal transformations of  $\mathbf{R}_{zz}$  of form

$$\mathbf{T} \mathbf{R}_{zz} \mathbf{T}^H = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{bmatrix} \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{T}_1^H & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2^H \end{bmatrix}, \quad (\text{A-13})$$

where  $\mathbf{T}_1 \in \mathbb{R}^{m \times m}$  and  $\mathbf{T}_2 \in \mathbb{R}^{n \times n}$  are nonsingular matrices [49]. This may easily be proved by showing that the coherence matrix of the transformed data  $\mathbf{T}_1 \mathbf{x}$  and  $\mathbf{T}_2 \mathbf{y}$  is the same as that of  $\mathbf{x}$  and  $\mathbf{y}$ .

In fact, the canonical correlations  $k_i$  form a *complete* or *maximal* set of invariants [49] for the composite covariance matrix  $\mathbf{R}_{zz} = E[\mathbf{z}\mathbf{z}^T]$ , under the linear transformation group

$$\mathcal{T} = \left\{ \mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 \end{bmatrix}, \det\{\mathbf{T}\} \neq 0 \right\}, \quad (\text{A-14})$$

with group action  $\mathbf{R}_{zz} \rightarrow \mathbf{T} \mathbf{R}_{zz} \mathbf{T}^T$  [49]. That is, any function of  $\mathbf{R}_{zz}$  that is invariant under the transformation group  $\mathcal{T}$  is a function of  $\mathbf{K}$ . This is the reason that the correlations  $k_i$  and coordinates  $\mathbf{u} = [u_i]_{i=1}^m$  and  $\mathbf{v} = [v_i]_{i=1}^n$  are called canonical [27].

## A.1 Linear Dependence and Coherence

The standard measure of linear dependence for the composite data vector  $\mathbf{z} = [\mathbf{x}^H \ \mathbf{y}^H]^H$  is the Hadamard ratio, inside the inequality

$$0 \leq \frac{\det\{R_{zz}\}}{\prod_{i=1}^{m+n} [R_{zz}]_{ii}} \leq 1, \quad (\text{A-15})$$

where  $[R_{zz}]_{ii}$ 's,  $i \in [1, m+n]$  are the diagonal elements of  $R_{zz}$ . This ratio takes the value 0 iff there is linear dependence among elements of  $\mathbf{z}$ ; it takes the value 1 iff the elements of  $\mathbf{z}$  are mutually uncorrelated.

By introducing a block Cholesky factorization [27, 51] for  $R_{zz}$  of the form

$$R_{zz} = \begin{bmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & R_{xy}R_{yy}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} Q_{xx} & \mathbf{0} \\ \mathbf{0} & R_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ R_{yy}^{-1}R_{yx} & \mathbf{I} \end{bmatrix}, \quad (\text{A-16})$$

where  $Q_{xx} = R_{xx} - R_{xy}R_{yy}^{-1}R_{yx}$  and is known as the error covariance matrix. It is then possible to write  $\det\{R_{zz}\}$  as

$$\begin{aligned} \det\{R_{zz}\} &= \det\{Q_{xx}\} \det\{R_{yy}\} \\ &= \det\{R_{xx}\} \frac{\det\{Q_{xx}\}}{\det\{R_{xx}\}} \det\{R_{yy}\}, \end{aligned} \quad (\text{A-17})$$

yielding the following decomposition of the Hadamard ratio:

$$\frac{\det\{R_{zz}\}}{\prod_{i=1}^{m+n} [R_{zz}]_{ii}} = \frac{\det\{R_{xx}\}}{\prod_{i=1}^m [R_{xx}]_{ii}} \det\{\mathbf{I} - \mathbf{K}\mathbf{K}^H\} \frac{\det\{R_{yy}\}}{\prod_{i=1}^n [R_{yy}]_{ii}}. \quad (\text{A-18})$$

The first and third terms on the right hand side of (A-18) measure the linear dependence *among* the elements of  $\mathbf{x}$  and  $\mathbf{y}$ , respectively, while the middle term,

$$L = \det(\mathbf{I} - \mathbf{K}\mathbf{K}^H) = \prod_{i=1}^n (1 - k_i^2); \quad 0 \leq L \leq 1, \quad (\text{A-19})$$

measures the linear dependence *between* the elements of  $\mathbf{x}$  and  $\mathbf{y}$ . The measure  $L$  takes the value 0 iff there is perfect linear dependence between elements of  $\mathbf{x}$  and  $\mathbf{y}$ ; it takes the value 1 iff the elements of  $\mathbf{x}$  and  $\mathbf{y}$  are independent. The  $i$ th term of the product on the right hand side of (A-19), i.e.  $(1 - k_i^2)$ , measures the linear

dependence between the  $i$ th canonical coordinate of  $\mathbf{x}$  and the  $i$ th canonical coordinate of  $\mathbf{y}$ . This implies that the linear dependence between  $\mathbf{x}$  and  $\mathbf{y}$  is decomposed into the linear dependence between their canonical coordinates, and is measured only by their canonical correlations or principal cosines.

Correspondingly, we may define the coherence measure between the elements of  $\mathbf{x}$  and  $\mathbf{y}$  as

$$H = 1 - L = 1 - \det(\mathbf{I} - \mathbf{K}\mathbf{K}^H) = 1 - \prod_{i=1}^n (1 - k_i^2); \quad 0 \leq H \leq 1. \quad (\text{A-20})$$

The elements of  $\mathbf{x}$  and  $\mathbf{y}$  are perfectly coherent iff  $H = 1$ ; and non-coherent iff  $H = 0$ .