

THESIS

BLOCK-BASED DETECTION METHODS FOR UNDERWATER TARGET
DETECTION AND CLASSIFICATION FROM ELECTRO-OPTICAL IMAGERY

Submitted by

Michael Jonathan Kabatek

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

Summer 2010

COLORADO STATE UNIVERSITY

June 1, 2010

WE HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER OUR SUPERVISION BY MICHAEL JONATHAN KABATEK ENTITLED BLOCK-BASED DETECTION METHODS FOR UNDERWATER TARGET DETECTION AND CLASSIFICATION FROM ELECTRO-OPTICAL IMAGERY BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE.

Committee on Graduate Work

Ali Pezeshki

Mingzhong Wu

Adviser: Mahmood R. Azimi-Sadjadi

Department Head: Anthony A. Maciejewski

ABSTRACT OF THESIS

BLOCK-BASED DETECTION METHODS FOR UNDERWATER TARGET DETECTION AND CLASSIFICATION FROM ELECTRO-OPTICAL IMAGERY

Detection and classification of underwater mine-like objects is a complicated problem due to various factors such as variations in the operating and environmental conditions, presence of spatially varying clutter, target obstruction and occlusion 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. Two different block-based methods are proposed for detecting frames and localization of mine-like objects from a new CCD-based Electro-optical (EO) imaging system. The block-based methods proposed in this study serve as an excellent tool for detection in low contrast frame sequences, as well as providing means for classifying detected objects as target or non-target objects. The detection methods employed provide frame location, automatic object segmentation, and accurate spatial locations of detected objects.

The problem studied in this work is the detection of mine-like objects from a new CCD imagery data set which consists of runs containing tens to hundreds of frames (taken by the CCD camera). The goal is to detect frames containing mine-like objects, as well as locating detected objects and segmenting them from the frame to be

subsequently classified as mine-like objects or background clutter. While object segmentation and classification of detected objects are also required as with the previous EO systems, the main challenge is successful frame detection with low false alarm rate. This has prompted research on new detection methods which utilize block-based snapshot information in order to identify potential frames containing targets, and spatially localize detected objects within those detected frames.

More specifically, we have addressed CCD object detection problem by developing block-based Gauss-Gauss and matched subspace formulations. The block-based detection framework is applied to raw CCD data directly from the sensor without the need for computationally expensive filtering or pre-processing as with the previous methods. The detector operates by measuring the log-likelihood ratio in each block of a given frame and provides a spatial 'likelihood map'. This detection process provides log-likelihood measurements of blocks in a given EO image which can then be thresholded to generate regions of interest within frame to be subsequently classified. This two-step process in both the Gauss-Gauss and matched subspace detectors consists of first measuring the log-likelihood, and determining frame of interest and then the regions of interest (ROI), and finally classifying the detected object ROIs, based upon shape-dependent features.

Complex Zernike moments are extracted from each region of interest which are subsequently used to classify detected objects. The shape-based Zernike moments provide rotational invariance, and robustness to noise which are desirable characteristics for classification. This block-based framework provides flexibility in the detection methods used prior to object classification, and solves the problem of having to invoke a classification system on every CCD frame by determining frames containing only potential targets.

A comprehensive study of the block-based detection and classification methods is carried out on a CCD imagery data set. A comparison is made on the detection and

false alarm rate performance for the Gauss-Gauss and matched subspace detectors on the CCD data sets acquired from the Applied Signal Technologies in Sunnyvale, CA. In addition a neural-network based classification system is employed to perform object classification based upon the extracted Zernike moments. The tested data set from AST consist of ten runs over the mine field each run containing up to several hundred frames. The total number of frames tested totals 1317, with 16 frames containing a single or partial targets in five of the data runs. Results illustrating the effectiveness of the proposed detection methods are presented in terms of correct detection and false alarm rates. It is observed that the low-rank Gauss-Gauss detector provides an overall frame detection rate of 100% at the cost of a false alarm rate of 36.9%. The matched subspace detector outperforms the Gauss-Gauss method and reduces the false frame detection rate by 16.9%. Using the Zernike features extracted from the matched subspace detector's output and an artificial neural network classifier yeilds a true frame detection rate of $P_d = 100\%$ at the cost of $P_{fd} = 16.8\%$ reducing the detected false frames detected by 3.3%. The reduced-rank Gauss-Gauss detector has a detection rate of $P_d = 100\%$ at the cost of probability of false detection $P_{fd} = 36.9\%$, using features extracted from the reduced-rank Gauss-Gauss detector's output passed to the neural network classifier yeilds a true detection rate of $P_d = 100\%$ at the cost of $P_{fd} = 21.7\%$ which significantly reduces the detected false frames by 15.1%.

Michael Jonathan Kabatek
Department of Electrical and Computer Engineering
Colorado State University
Fort Collins, CO 80523
Summer 2010

TABLE OF CONTENTS

SIGNATURE	ii
ABSTRACT OF THESIS	iii
1 INTRODUCTION	1
1.1 Problem Statement and Motivations	1
1.2 Literature Review	2
1.3 Objectives of the Present Research	8
1.4 Organization of the Thesis	9
2 EO SENSOR, DATA DESCRIPTION AND CHALLENGES	11
2.1 Introduction	11
2.2 CCD Sensor Description & Properties	11
2.3 CCD Sensor Data & Challenges	14
2.4 Conclusion	21
3 BLOCK-BASED GAUSS-GAUSS DETECTION	22
3.1 Introduction	22
3.2 Block-Based Detection	23
3.2.1 Review of Binary Hypothesis Testing	23
3.2.2 Rank Reduction	28

3.2.3	Implementation of Block-based Detection	31
3.2.4	Comparison Between Reduced-Rank and Full-Rank Detectors	33
3.2.5	Detector Design	33
3.3	ROI Segmentation	36
3.4	Conclusion	43
4	MATCHED SUBSPACE DETECTION	44
4.1	Introduction	44
4.2	Matched Subspace Detection	45
4.3	Comparison between Matched Subspace & Gauss-Gauss Detection .	49
4.4	Conclusion	54
5	FEATURE EXTRACTION IN ELECTRO-OPTICAL IMAGERY	56
5.1	Introduction	56
5.2	Shape-Dependent Feature Extraction	57
5.2.1	Zernike Moments	57
5.2.2	Feature Space Dimensionality Reduction	59
5.3	ROI Classification	60
5.4	Classification Results and Observations	63
5.4.1	ROI Classification Results	63
5.4.2	Overall FOI Detection/Classification Results	66

5.5	Conclusion	73
6	CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK	75
6.1	Future Work	77
APPENDIX A	— GRAPHICAL USER INTERFACE (GUI) . . .	85

LIST OF FIGURES

2.1	Photo of CCD sensor courtesy of Richard Manley, NSW-PCD.	13
2.2	Frame sequence example from run TargetY8_001 containing target frames (full and partial target frames are shown).	17
2.3	Typical Target and Non-Target Frames.	18
2.4	Typical Target and Non-Target histograms.	18
2.5	All targets contained in Table 2.1.	19
2.6	Selected non-targets clutter contained in Table 2.1.	20
3.1	Block-based detection process.	31
3.2	Comparison between full rank and reduced rank detectors.	34
3.3	FOI from data run SAM001_004 used to train the Gauss-Gauss detector. Regions of blocks were selected over the target and over background.	35
3.4	Detection and ROI segmentation process.	38
3.5	Various detector outputs for different target frames.	40
3.6	Various detector outputs for different detected background anomalies.	41
4.1	Covariance structure computed using \mathbf{N} (left) and \mathbf{T} (right) data matrices. These matrices show large values along the diagonal and lower values off the diagonal.	48

4.2	Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run SAM23_005 frame 162	50
4.3	Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run TargetY8_001 frame 71	51
4.4	Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run SAM23_004 frame 156	52
4.5	ROC for FOI performance comparison of the reduced-rank Gauss-Gauss and matched subspace detectors.	54
5.1	Several target and non-target ROI's used for the BPNN training set.	59
5.2	ROI classification performance receiver operating characteristics (ROC) for matched subspace detector, and reduced-rank Gauss-Gauss detector.	65
5.3	FOI performance receiver operating characteristics (ROC), with and without BPNN.	68
5.4	Plot of mean and standard deviation for Zernike features in the reduced feature sets	70
5.5	Example of false negative (ROI's). The likelihood map output from matched subspace detector (left), and false negative ROIs (right). . .	71

5.6	Example of false positive (ROI's). The likelihood map output from matched subspace detector (left), and false negative ROIs (right). . .	72
A-1	GUI tool used for building training sets for the detector, and feature sets for the classifier.	86

LIST OF TABLES

2.1	Tested CCD Data Set	14
3.1	Detector output Measures	39
3.2	Detection Results for Reduced-Rank Gauss-Gauss	42
4.1	Detection Results for Matched Subspace Detector	53
4.2	Detection Results for Reduced-Rank Gauss-Gauss Detector	53
5.1	MATLAB Neural Network Training Parameters	62
5.2	Confusion Matrix for Matched Subspace Output	64
5.3	Confusion Matrix for Reduced-Rank Gauss-Gauss Output	64
5.4	Detection Results for Matched Subspace Detection and Classification	67
5.5	Detection Results for Reduced-Rank Gauss-Gauss Detection and Clas- sification	67
A.1	File Menu	89
A.2	GUI Plots and Graphical Data Display	90
A.3	GUI Controls	91

CHAPTER 1

INTRODUCTION

1.1 Problem Statement and Motivations

An underwater mine is a self-contained explosive device placed in water to destroy ships or submarines. Ocean mines have been a major threat to the safety of vessels and human lives for many years. The Navy's capability to conduct shallow water and very shallow water mine countermeasures in support of beach assaults, as well as keeping the ocean as a safe place is a very important issue, and current mine-hunting technology is still in need of major improvement [1] - [12]. To clear the threat of the naval mines and ensure that the fleet can carry out operations in the open ocean and littoral, including maintaining open sea lanes of communication and supporting maneuver warfare from the sea, the US Navy has devoted substantial resources and efforts to detect and discriminate different types of underwater mines. In order to improve the Navy's ability to effectively prevent other nations from posing a significant threat to the national security or economy of the US by mining in the oceans, extensive research and developmental work on underwater mine detection, classification, and identification have been supported for many years [1] - [12].

The problem of detection of underwater objects in electro optical imagery has been mainly carried out using two typed of EO sensors. The sensors used for underwater mine detection include the laser line scan (LLS) technology, and Streak Tube Imaging LIDAR (STIL). These systems although different operate similarly by scanning line-by-line over a target field in order to identify potential mine like objects arising from sonar contact. The LLS and STIL imaging systems generate two dimensional contrast and range data: The bottom return includes both time of flight information, which

provides a quantitative measure of the height of the object above the bottom and the radiometric level that is proportional to the reflectivity of the bottom object. In contrast to the new CCD-based imaging system which provides sequences of contrast images (photographs) over the target field.

Although the sensor technology for underwater mine identification has advanced to a level that these systems are being transitioned into the fleet, the target identification is still being done by human operators [1]- [12]. The development of an automatic underwater target identification system capable of identifying various types of underwater targets (mines), under different environmental conditions pose many technical problems. Some of the contributing factors are: Targets have diverse sizes, shapes and reflectivity properties; target emplacement environment is variable; targets may be proud or partially buried; environmental properties vary significantly from one location to another. Bottom features such as sand, rocks, corals, and vegetation can conceal a target whether it is partially buried or proud. Competing clutter with responses that closely resemble those of the targets may lead to significant number of false positives. All these factors contribute to make this problem a very complicated and challenging one.

1.2 Literature Review

Identification of mine-like objects is a pressing need for military, and other ocean fleets. In mine countermeasures operations, sonar is used to detect and classify mine-like objects if their sonar signatures are sufficiently similar to known signatures of mines. For littoral regions, its possible that hundreds of mine-like objects need to be identified for safe passage of the Fleet [12]. This operation is a time-consuming identification process performed manually by Explosive Ordnance Disposal divers or Remotely Operated Vehicles. Rapid visual identification of mine-like objects using electro-optic identification sensors can dramatically improve the time required for

mine countermeasures operations.

Electro-optical (EO) imaging systems [1] - [12] are being increasingly exploited as target identification tools with good spatial resolution. To support rapid visual identification, two types of electro-optic identification (EOID) sensors have been under investigation by the Navy. These laser identification systems used are: the Arete Associates Streak Tube Imaging LIDAR (STIL) system, the Northrop Grumman Laser Line Scan (LLS) system, and the Raytheon LLS system [12]. In [12] the two main EO sensors are described. The EOID laser line scan technology uses a diode-pumped Nd: YAG laser that provides 500 mW (Raytheon system) and 160 mW (Northrop Grumman system) of power, both operating at 532 nm wavelength. The laser illuminates a small spot, which is synchronously scanned by a photo-multiplier receiver to build up a raster-scanned image. The laser scans downward through a 70-degree field-of-view [12].

Arete Associates developed the patented STIL technology specifically for high-resolution three-dimensional imaging of underwater objects. The STIL system is an active imaging system using a pulsed laser transmitter and a streak tube receiver to time resolve the returned light. The laser beam is diverged in one dimension using a cylindrical lens to form a fan beam. The returned light is imaged onto a slit in front of the streak tube photocathode by a conventional lens, and is time (range) resolved by electrostatic sweep within the streak tube, generating a 2-D range-azimuth image on each laser pulse. The bottom return includes both time of flight information, which provides a quantitative measure of the height of the object above the bottom and the radiometric level that is proportional to the reflectivity of the bottom object. Each laser shot thus provides range to and contrast of the bottom for each cross-track pixel [12].

The work in [3] overviews the EOID sensors project for developing a Laser Visual Identification Sensor (LVIS) for identification of proud, partially buried, and moored

mines in shallow water and very shallow water, which is deployed in small diameter underwater vehicles, including unmanned underwater vehicles (UUVs). The authors in [3] state that LVIS must: a) deliver high quality images in turbid coastal waters, while b) being compatible with the size and power constraints imposed by the intended deployment platforms. LVIS is designed to produce images of mine like contacts (MLC) of sufficient quality to allow identification while operating in turbid coastal waters from a small diameter UUV.

Technology goals in [3] are: a) identification range up to 40 feet for proud, partially buried, and moored MLCs under coastal water conditions; b) day/night operation from a UUV operating at speeds up to 4 knots; c) power consumption less than 500 watts, with 275 watts being typical; and d) packaged within a 32-inch long portion of a 21-inch diameter vehicle section.

The work in [6] described various spatial and non-spatial sensing concepts and discuss hardware implementations. In particular they highlight the ability of laser-based systems to produce imagery at very low light levels. The authors in [6] present the utility of low light level imagery in two dimensions, the potential benefits of three-dimensional low light level imagery as well as characteristics of systems that can implement these concepts.

Previous work [1] - [12] on the development of detection and classification methods have been focused on the data sets collected using the two described EO sensors [12]. The work in [1] used the data collected using the STIL which produced high-resolution 3-D images of underwater objects. STIL scans line by line, on a rectangular area of a target field [1]. The collected raw STIL data is rendered to produce pairs of contrast (gray-level) and range (distance) maps [5]- [12]. The previous work [1] - [4] focused on filtering, segmentation, and classification of underwater mine-like objects from cropped regions of the STIL scans.

In [1] three filtering methods were tried to preprocess the STIL images. Each

cropped STIL image containing a mine-like object was filtered using three filters in succession: a median filter, k-nearest mean (KNM) filter, and an edge preserving filter [1]. These filters attempt to remove background noise, and sharpen edges before the segmentation stage. After preprocessing the mine-like objects must be segmented from the STIL images. Two different segmentation methods namely a global-based histogram modeling, and a contour-based method were studied in [1]. In this work histogram modeling estimates the background parameters, and uses a maximum likelihood (ML) -based method for removing the background. In histogram-based background/noise removal methods, different PDF's such as Gaussian, Rayleigh, Gamma, uniform, exponential, and Bernoulli were tried to model the modes in the histograms. The global histogram modeling process used in [1] assumed that the signal and noise are comprised of a two-component Gaussian mixture. The parameters of these two Gaussian's were found using the expectation maximization (EM) iterative algorithm [22]. From this process a threshold is found to segment the background from the mine-like object in the STIL image. The second object segmentation method explored in [1] and [4] was a contour-based method using gradient vector flow (GVF) snake [23]. Using this method an initial contour is set that can move under the influence of internal force parameters from within the curve itself and external forces computed from the image data. The internal and external force parameters are defined so that the snake will conform to an object boundary or other desired features within an image. Canny edge map [24] is initially used to detect the edges of the mine-like objects, then an initial contour is set, and deformed until convergence is achieved.

Once object silhouettes are generated for both contrast and range STIL images features are extracted from the combined silhouettes. The features extracted from the segmented images included Zernike moment shape dependent features [25], [26], and

Gray Level Co-occurrence Matrices (GLCM) texture-based features [18] which are extracted from both range and contrast maps within the silhouette boundary. Zernike moments are shape dependent features based on Zernike polynomials. GLCM, on the other hand, computes several statistical/textural features namely contrast, correlation, entropy, and homogeneity. Various feature extraction schemes are available that can be used to extract shape-dependent features for a wide variety of pattern recognition problems. However, moment-based schemes [27]- [28] are among the most widely used methods as they provide translation, rotation and scaling invariant features ideal for 2-D as well as 3-D pattern recognition applications. In [30], a comparison is made among several types of moments including regular moments, Legendre moments, Zernike moments, and complex moments. These methods were compared in terms of their image representation ability, noise sensitivity, and information redundancy on several character recognition examples. Owing to the fact that the regular moments do not provide an orthogonal representation, the extracted features using this scheme lack optimality in representation. This is in contrast to the orthogonal moments, e.g. Legendre and Zernike moments [27]- [28]. The experiments conducted in [29] indicated that the classification results of the Zernike moments are substantially less sensitive to additive noise effects in the images when compared to the other types.

In [28], a similar study was carried out where the regular moments and Zernike moments were used for subsequent feature extraction and a back-propagation neural network (BPNN) [19] was employed as a classifier. The system was tested for classifying 26 uppercase characters (A to Z) in the English alphabets. The silhouettes were allowed to have varying scale, translation and orientation forming 24 sets of images. In addition, random noise with varying SNR from 5 to 50 dB was added to the patterns. The simulation results once again showed the noise immunity of the Zernike moments particularly when used in conjunction with a BPNN classifier. In another

study [27] Zernike moments were used for recognition and pose estimation of 3-D objects from the 2-D perspective views. The scheme utilizes multiple BPNN's with different parameters and structures. The decisions of these networks were fused together using a majority voting scheme. It was observed that combining the decisions of these parallel networks can minimize the occurrence of erroneous decisions. Due to the use of Zernike moments the performance was invariant to viewing angle, location and orientation of the objects in the image. The effectiveness of the system was demonstrated on several clean and noisy patterns of military ground targets. Finally, the two pose parameters, namely elevation and aspect angles, were estimated using a two-stage neural network structure. In [31], a pattern classification scheme for classifying buried land mines of wood and nylon compositions from microwave imagery data was developed. Two-dimensional (2-D) Karhonen Loeve (KL) transform and Zernike moments were used to extract energy and shape-dependent features of the segmented land mine regions. A neural network was then trained to discriminate the targets from the non-target anomalies. The comparison of the results indicated that the Zernike moments gave much better discrimination of wooden type mines that are generally very difficult to identify due to their weak response in the microwave images. This is due to the property that the dielectric constant of wood is closer to that of soil than the nylon. Additionally, it was observed that the uncorrelated property of these feature extraction schemes substantially improved the training of the neural network classifier. Experimental results in [1] show Zernike moments remained robust, and invariant to rotation, and scaling, while they changed for different grazing angles. The GLCM texture based features proved to be more robust to grazing angle changes. Both Zernike moments, and GLCM were used as features in the classification process. Among classifiers used are back-propagation neural network [19] and support vector machines [20] were tried to classify mine-like objects in STIL images. Due to all the useful properties of the Zernike moments we have chosen to use this

method for shape-dependent feature extraction of detected regions of interest (ROI) in our study.

1.3 Objectives of the Present Research

The work in this research project used a different EO sensor consisting of a CCD camera and LED illuminator. Due to the differences between the EO CCD and STIL systems a new detection framework was determined to be necessary. The key difference between the CCD system and the STIL system is the way they operate and generate images. The STIL sensor scans a target field line-by-line in order to acquire an image containing mine-like objects, whereas the CCD system takes a sequence of snapshots over a target field as the vehicle carrying the sensors is moving through the target field. For the previous work the STIL images (both range and contrast) were cropped by hand, preprocessed, segmented, and classified in order to determine if the detected object was a target or non-target as well as identifying different types of mines. For the CCD case, a data run is generated which contains many snapshots (30 to 300 frames) of a target field. The majority of the frames in a CCD data run contain only background clutter or partial targets. Therefore, it is necessary to not only segment and classify mine-like objects (as with STIL), but also to automatically determine the frame(s) of interest (FOI) containing mine-like objects within a data run. This automatic determination of the FOI for a given data run is necessary otherwise hundreds of frames would need to be preprocessed, segmented, and passed to a classification system, resulting in tremendous computational overhead. The added complexity of automatically determining FOI from a data run make this problem a very challenging one.

After investigating many possible target detection and segmentation schemes, it was decided to develop a low-rank block-based Gauss-Gauss [16, 14], as well as a

matched subspace detectors [15] to resolve both the problem of automatic determination of FOI within a data run, as well as automatic object segmentation for subsequent feature extraction. The proposed method for detection of mine-like objects in the EO CCD database involves a local block-based detection in the spatial domain. Each CCD contrast image in a data run is partitioned into 4×4 blocks. Each block is then applied to the detector in order to determine if the block contains background or a portion of a potential mine-like object. If the object passes the criterion of being a mine-like object it is flagged as a target block. If a predetermined number of connected blocks in an image are determined to be target blocks, the whole frame is flagged as FOI. The connected target blocks detected in the FOI comprise the ROI allowing automatic segmentation for feature extraction. The connected target blocks now comprise a segmented image of a potential target. From this segmented image map, features can easily be extracted and passed to a classifier to determine if the mine-like object is a target or non-target. Results show that FOI identification and ROI classification can be achieved for all targets in the studied data set. Comparing to the work in [1] our overall system is substantially less complicated since no separate preprocessing and segmentation is needed. This makes adoption in real mine hunting systems that use EO sensors a reality.

1.4 Organization of the Thesis

The organization of this thesis is as follows. Chapter 2 describes characteristics of the CCD sensor; and data collection methods as well as a description of the CCD imagery data for targets (mines) and non-targets (background clutter). The chapter also presents the data set, and challenges associated with detection on CCD frame sequences. In Chapter 3 a review of binary hypothesis testing and Gauss-Gauss detection is presented. Next rank reduction is presented for Gauss-Gauss detection followed by the implementation of block-based detection algorithms on the CCD data

set. We also present a comparison of full-rank and reduced-rank Gauss-Gauss detectors on the tested data set, as well as details of the process which is used to design the detectors. The process involved to detect FOI and segment ROI using the Gauss-Gauss detector is also presented in 3. In Chapter 4, we present the matched subspace detector and relevant theory. We compare Gauss-Gauss detection and matched subspace methods and show how the latter detector improves overall clutter suppression. In Chapter 5 we present results on feature extraction methods, as well as results of the target classification using shape based Zernike features. We present a comparison of the classifiers applied to both the results of the Gauss-Gauss detector as well as those of the matched subspace detector. Chapter 6 concludes the studies carried out in this research and discusses the goals for future work.

CHAPTER 2

EO SENSOR, DATA DESCRIPTION AND CHALLENGES

2.1 Introduction

In this chapter we present several aspects of the data collection process involved with the EO-CCD imaging system including properties of the sensor, details on how the data is collected, and description of the challenges associated with detection in EO-CCD images. Since the data collection using the EO-CCD system involves capturing a sequence of frames over a target field, which is in contrast to scanning line-by-line over a target field (as in the previous data from the STIL system [1]- [12]), a new framework for detection had to be developed to work on this new class of images.

This chapter is organized as follows: in Section 2.2 we first present the technical aspects of the CCD sensor including the CCD camera used, camera resolution, and other information about the imaging system used in this research. Next, in Section 2.3 a discussion of the data produced by the CCD sensor is presented, and information about the tested data set is reviewed. Finally, we discuss challenges involved in automatic target recognition of mine-like objects using the CCD system, and also present examples of frames contained within the CCD data set, followed by concluding remarks about this CCD imagery data set.

2.2 CCD Sensor Description & Properties

This section provides a detailed overview of the EO-CCD module. The EO system developed by the Applied Signal Technologies, Inc (AST), in conjunction with the

Naval Surface Warfare Center, Panama City (NSWC-PCD) consists of four main components, which are: the 12-bit CCD camera, PC104 stack, on-board & external hard drive, and external changeable LED illuminators [13]. In the following we will describe the EO module CCD system, and its components.

The EO module (shown in Figure 2.1) employs a DVC-1500M monochrome CCD camera. The camera is used to take ocean bottom photos (frames) over a target field. The CCD camera is a high performance digital camera with functions tailored to high-throughput scientific and industrial applications. It is capable of both high-speed readout (40 MHz pixel rate) and low noise readout (20 MHz pixel rate) at 12 bits [32]. It utilizes a Sony ICX285AL progressive scan interline CCD. The high quantum efficiency of the CCD peaks in the 500-600 nm region of the spectrum [32]. The CCD camera has four basic operating modes: streaming overlapped exposure, streaming non-overlapped exposure, edge-triggered single frame snapshot, and variable pulse-width exposure. Each mode can be operated at either 20 or 40 MHz and can support variable binning and region of interest operation. The camera is capable of producing images having sizes of 1394×1040 (6.45um pixel size) having a high dynamic range of 12-bit/pixel with multiple binning modes (1×1 to 8×8) [13]. The next main component in the system is the PC104 stack. The PC104 stack is the main computer on the EO module which controls all subsystems including the CCD camera, LED illuminator, and storage devices. The PC104 stack consists of a 1GZ Pentium 4 with 1GB Ram, 1394b firewire II board and Ethernet switch board. The next main component are the storage devices for storing captured images from the camera (frames). The external and internal storage consist of one internal 100GB hard drive and one external 100GB firewire hard drive. The final system component is the LED illuminator used to provide more light for the CCD camera when capturing ocean bottom photos. The LED illuminator on the EO module is a white Philips Lumiled Luxeon Flood 18 LED illuminator. The LED illuminator consists of 18 Indium

Gallium Nitride (InGaN) LED light sources mounted onto an aluminum core printed circuit board and provides accurate light center positioning with Luminous Flux > 500 lumens [33]. The EO module is capable of active or passive imagery (illuminator on or illuminator off while capturing ocean bottom photos), and is designed to operate over a large dynamic range to maximize imaging capability in turbid water conditions [13].

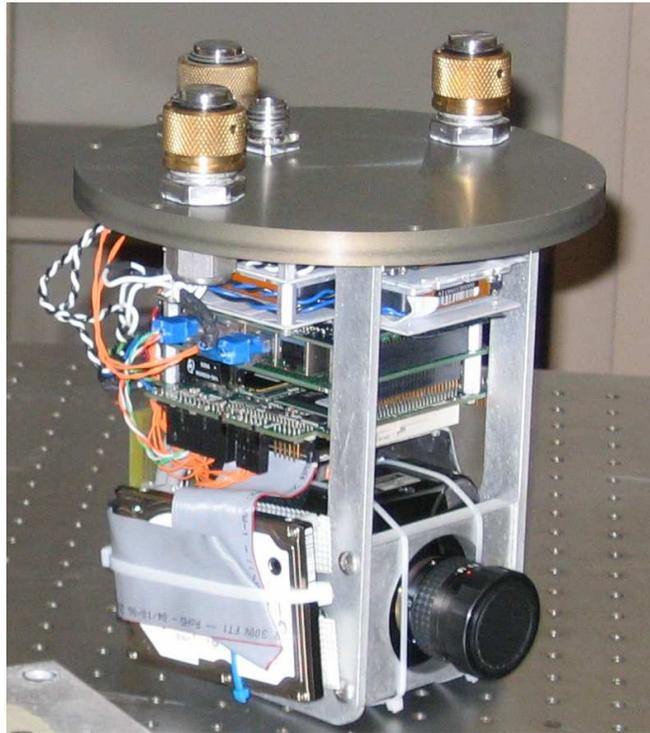


Figure 2.1: Photo of CCD sensor courtesy of Richard Manley, NSW-PCD.

The EO module has a length of 8.5 in. and a diameter of 8 in., it weights 15 lbs, and has a payload size of 12 in. [13]. The EO module is on a Bluefin-12 autonomous underwater vehicle (AUV) developed by Bluefin Robotics [34]. The Bluefin-12 has variable payload flexibility and capability the Bluefin-12 is light weight, and is tailored to support a wide range of payloads in the forward 48 of the vehicle. The bluefin-12

has low self-noise, outstanding dynamic control, and magnetic and inertial navigation providing payload data quality [34]. The vehicle has a flooded architecture, acoustically transparent shell material. The EO module is housed in the vehicle which takes sequential photos (frames) of the ocean bottom in order to detect proud and/or buried mines. An example frame sequence is shown in Figure 2.2.

2.3 CCD Sensor Data & Challenges

The CCD image data consists of a series of ocean bottom snapshots as can be seen in Figure 2.2. The data analyzed consists of five data runs containing targets, and five containing no targets (just background). The data runs used as a testing set in this study together with total number of frames per run and target FOI's are given in Table 2.1.

Table 2.1: Tested CCD Data Set

Run	Total Frames	FOI
SAM001_003	42	0
SAM004_001	35	0
SAM22_011	35	0
SAM23_003	293	3
SAM23_004	287	3
SAM23_005	293	4
TargetY8_001	136	3
TargetY8_003	29	0
TargetY8_004	32	0
TargetY8_006	135	3
Totals	1317	16

In this study a total number of 10 data runs were analyzed. The total number of frames in the data runs is 1317, of which 16 frames contain targets. The CCD system produces images that are 684×513 pixels at 12-bits per pixel gray level resolution. The data imported is resized to 512×512 pixels for ease of computation using the default MATLAB bicubic interpolation image resizing algorithm. Example

data frames containing target (mine) and non-target (background only) are shown in Figure 2.3 which exhibit considerable overlap. The histograms for these typical target and background only (non-target) frames are shown in Figure 2.4. The data set contains only two different types of targets: runs TargetY8_001, TargetY8_006 contain long cylindrical targets as found in Figure 2.3; where as runs SAM23_003, SAM23_004, and SAM23_005 contain partial targets of different shapes as shown in Figure 2.5. Figure 2.6 shows several non-target frames with varying bottom conditions.

There are three main challenges involved in designing an automatic target detection and recognition system for the new EO database. The first is FOI detection, which is the key to the success of other subsequent steps namely feature extraction and classification. Since only a few out of several hundred frames in a run may contain a partial or full target images, it is important to isolate only those frames which contain a potential target. The next challenge is segmentation of the mine-like objects within the FOI for ROI selection. This is another main challenge due to the fact that background and mine-like objects tend to have very similar contrast and texture characteristics, hence making the segmentation and discrimination very difficult tasks. The third challenge involved with designing robust target detection and classification systems for this new CCD EO database lies in the fact that FOI may contain partial targets. Partial targets cause difficulties for both detection and classification systems due to the fact that the extracted ROI may not contain adequate discriminatory information. These challenges and issues are discussed in more detail below.

1. As mentioned before each data run contains such a large number of frames containing only background and few frames containing targets. The focus of this work is to detect FOI within the runs, and extract ROI only from the detected frames containing potential targets. Once ROI are extracted from the FOI, the problem becomes a two-class classification problem to determine if

the detected object is a target or a non-target. The main challenge involves designing a detector that will provide screening mechanism to filter out frames that have no object of interest. If a mine-like object exists in a frame, the frame must be marked as FOI, so that the detected objects contained in the FOI can be segmented and further classified.

2. The next main challenge when designing a detection system for this new EO database is successful ROI segmentation. We can see from our typical target and background frames in Figure 2.3 and their distributions in Figure 2.4 that the background and target have overlapping gray level intensities. This makes it difficult to employ global-based schemes to segment the detected objects. Also low contrast of the CCD EO images do not provide any identifiable texture to discriminate between target and background.
3. Finally partial targets are fragmented ROI within a FOI (See Figure 2.5). This can occur because of occlusion or when only a portion of a mine-like object is captured in a frame, hence causing two problems. The first problem is the fact that a partial target may be very small and indistinguishable from background anomalies (see Figure 2.6). Small objects pose a challenge since the detector must have some way of discriminating small anomalies from very small portions of targets. Another issue involved in partial targets exists in the fact that these small ROI must be classified after they are detected. A classifier may incorrectly classify a partial target due to the lack of adequate discriminatory features.

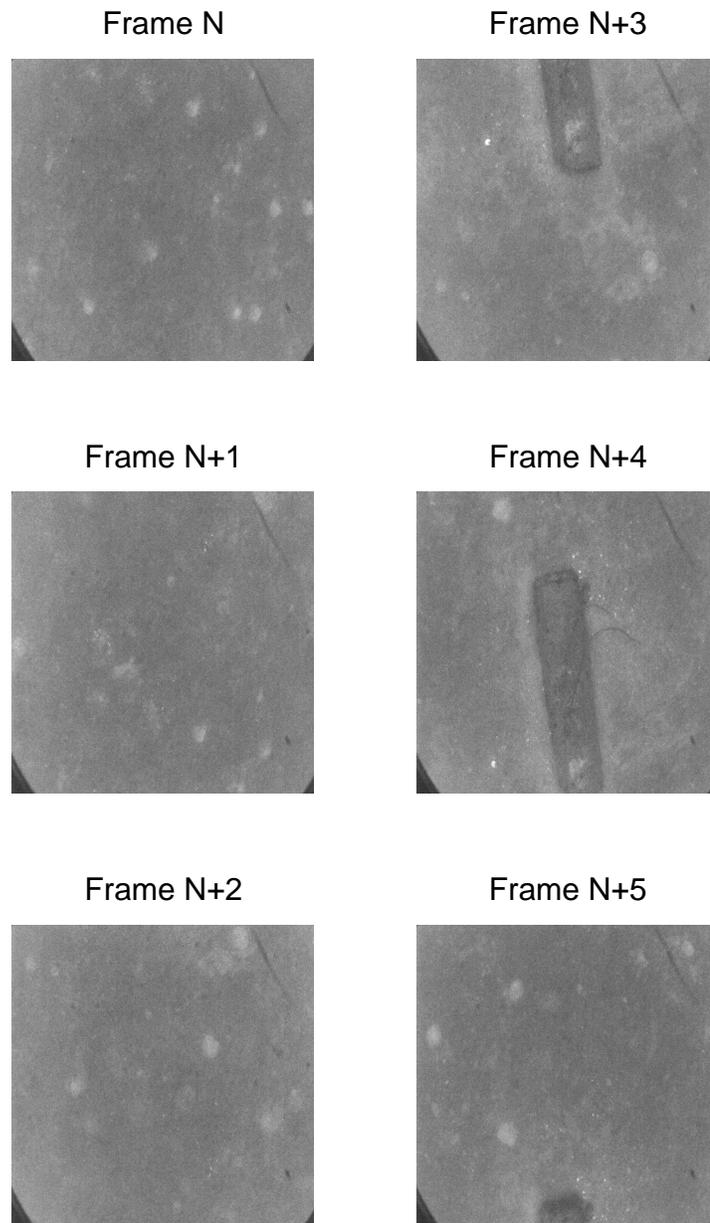


Figure 2.2: Frame sequence example from run TargetY8_001 containing target frames (full and partial target frames are shown).

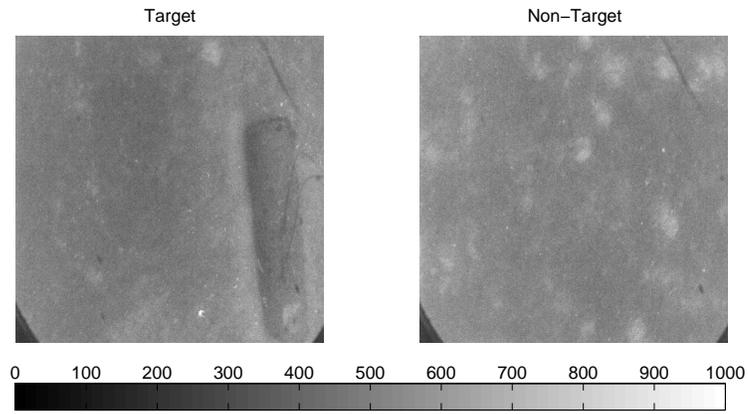


Figure 2.3: Typical Target and Non-Target Frames.

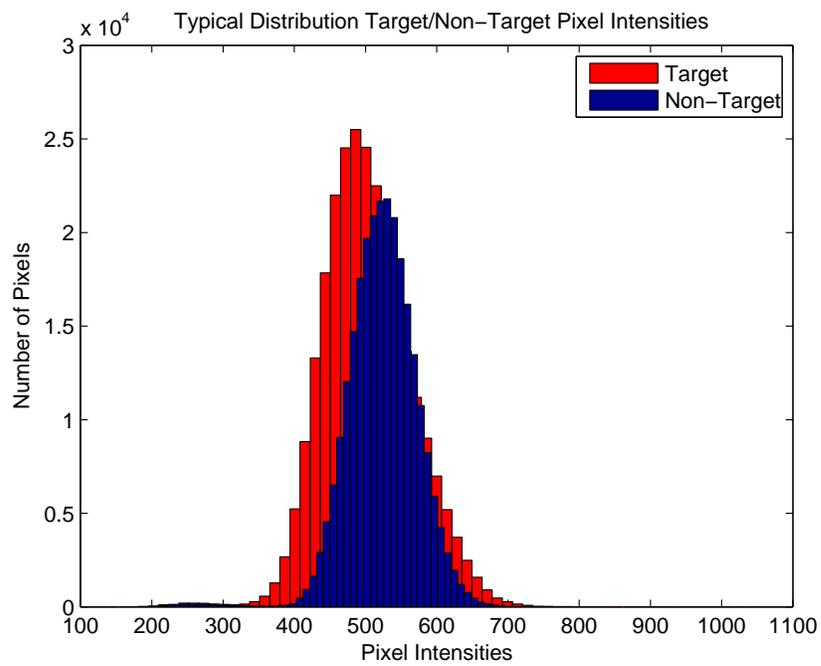


Figure 2.4: Typical Target and Non-Target histograms.

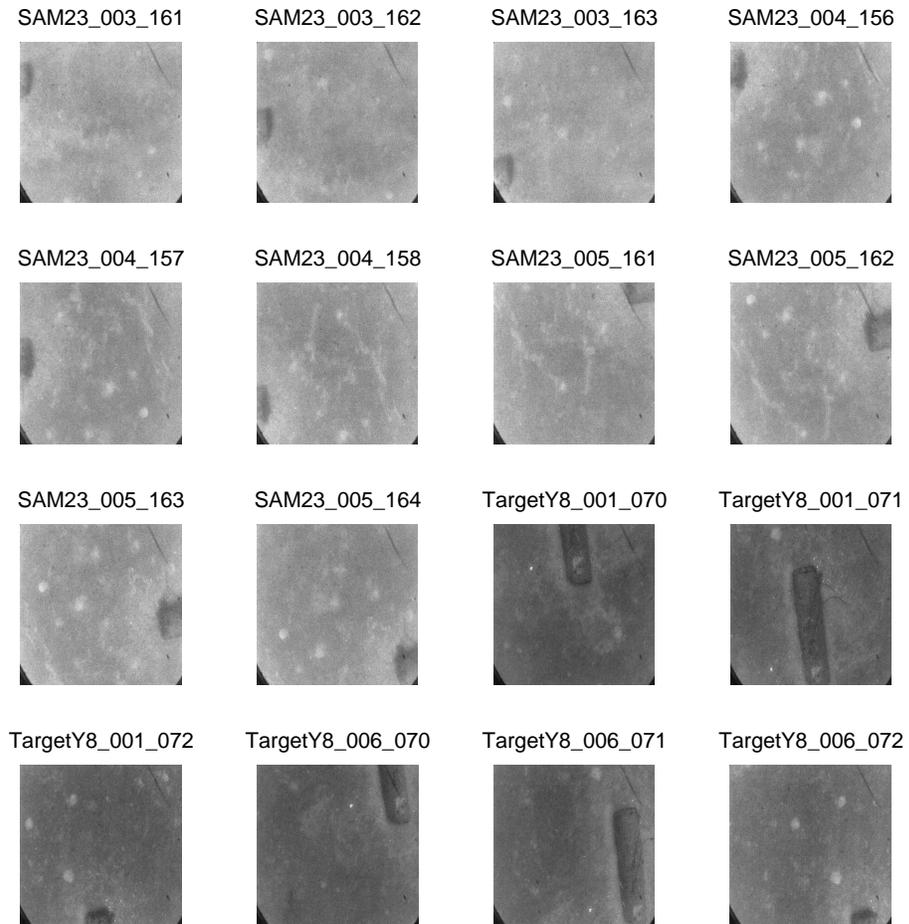


Figure 2.5: All targets contained in Table 2.1.

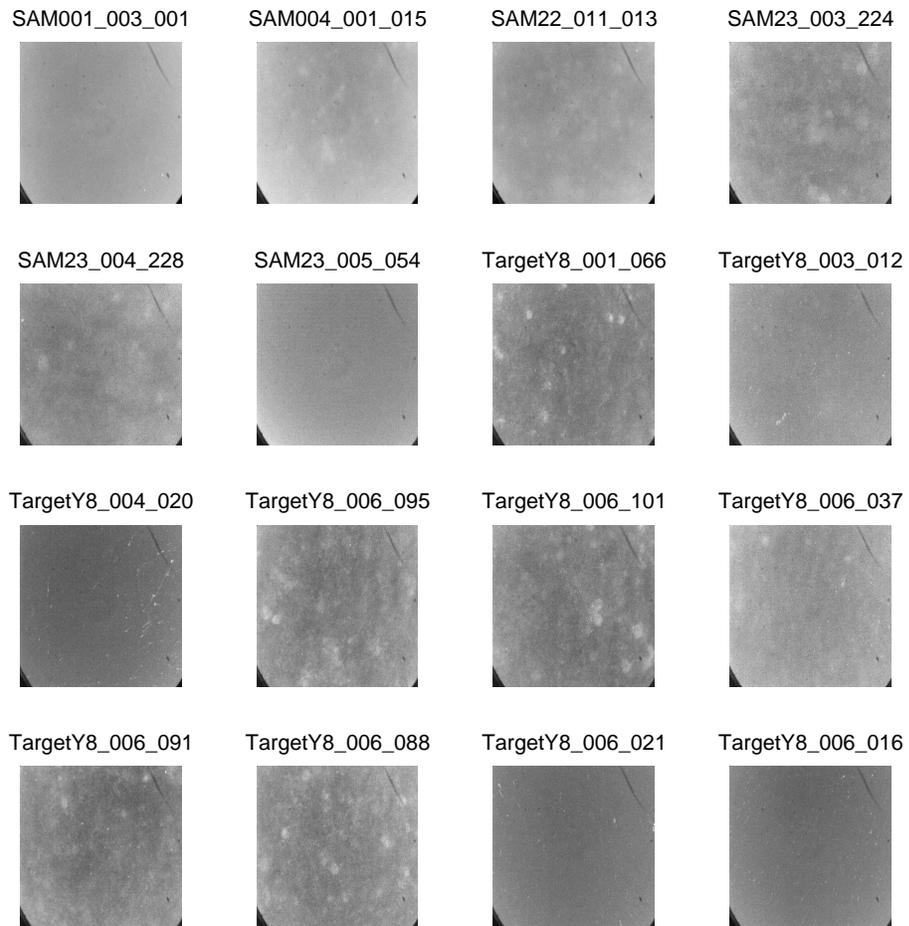


Figure 2.6: Selected non-targets clutter contained in Table 2.1.

2.4 Conclusion

In this chapter we presented the description and properties of the CCD sensor used for collecting the EO images, the type of data produced from the sensor, and those used in this study, as well as challenges associated with designing an automatic target detection and recognition system for the data collected using this new sensor. The sensor provides data runs which consist of sequences of ocean bottom snapshots in which FOI must be first detected. Subsequently ROIs that contain potential mine-like objects must be segmented in order to extract salient shape dependent features to classify them as mine-like objects or background anomalies. The main challenges in this work are: (1) FOI detection, (2) ROI segmentation, and (3) partial target feature extraction. In contrast to the STIL sensor that produced a pair of contrast and range images this CCD sensor produces only one image with typically poor contrast between target objects and background regions. These overlapping pixel intensities make it difficult to apply global-based schemes over the entire image to segment the potential targets. For these reasons a block-based scheme is employed for target detection and segmentation. This is discussed in the next chapter.

CHAPTER 3

BLOCK-BASED GAUSS-GAUSS DETECTION

3.1 Introduction

In this chapter a block-based method for detection of FOI within a run, and determining ROI within the detected frames is described. The main reason for taking a local-based (block) approach as opposed to a global-based approach employed on the STIL data [1, 2, 4] lies in the fact that FOI must be determined for every data run. If a histogram (global-based) approach were to be employed here preprocessing and segmentation would be performed on every frame in the data set. However, as mentioned before in the CCD-based database mine-like objects tend to have the same pixel intensity as those of background regions, hence making global-based methods inefficient for this application. In contrast, in the local block-based approach each image is processed block-by-block using a local-based Gauss-Gauss detector [14, 16] which exploits local statistical (second order) properties of the mine-like objects, and background anomalies. Only blocks that have similar characteristics to mine-like objects are flagged as detections. Once all blocks within a given frame are processed then a collection of connected blocks will be defined. Conceivably this method should identify all the blocks in a given frame that belong to a potential mine-like object. This collection of connected blocks will result in a segmented mine-like object from which features will be extracted. The proposed local-based method accomplishes two goals: (a) determines if an object (or part of an object) exists in a frame thereby detecting a FOI. This reduces the number of frames which need to be looked at by the classifier; and (b) automatically gives the location of the potential mine-like object, and segments the ROI with mine-like characteristics from the FOI. In what follows

we describe the theory and results of this local-based detector and its reduced rank version.

In this chapter we review binary hypothesis testing in Section 3.2 as well as the measurement model, and details regarding the model used in this study. Section 3.2 also presents the Gauss-Gauss formulation of the detector, and methods for detection improvements using rank reduction [14]. Details on how the block-based detector is implemented on the EO-CCD data and description of the procedures involved with generating detection measures from the EO-CCD images and generating FOIs are also presented. This section presents a comparison between full-rank, and reduced-rank detection, and describes procedures involved in the detector design. Section 3.3 of this chapter explains in details how ROIs are segmented from detected FOIs and presents several examples of detected ROIs as well as the detector’s performance on the tested EO-CCD data set in Table 2.1.

3.2 Block-Based Detection

To determine FOI the sequence of frames in a data run are partitioned into small blocks of size 4x4 and the problem is casted as block-based binary hypothesis testing. A brief review of binary hypothesis testing using Neyman-Pearson and Gauss-Gauss [14, 16] detection is given in the next subsection. In Chapter 4 we present an improved version of the detector that uses the matched subspace method [15, 17].

3.2.1 Review of Binary Hypothesis Testing

The classical detection problem of choosing between two hypotheses [16] is that given an N-dimensional observation space, where $\mathbf{y} = [y_1, y_2, \dots, y_N]^H$ represents an observation (measurement) vector in this space, we would like to test between H_1 hypothesis (true) and H_0 hypothesis (null) for this observation vector. In this specific problem our observations (or measurements) are pixel blocks of $n \times n$ ($n = 4$) pixels shaped into column vectors of size $N \times 1$ therefore $N = n^2$, and for this detection

problem under H_1 our measurement \mathbf{y} contains signal plus noise, while under H_0 our measurement contains noise alone. That is,

$$\begin{aligned} H_1 & : \mathbf{y} = \mathbf{x} + \mathbf{n} \\ H_0 & : \mathbf{y} = \mathbf{n} \end{aligned}$$

where \mathbf{x} represents the signal and \mathbf{n} represents the noise. Clearly, each time we conduct the test there are four possible outcomes. These 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 are denoted by, C_{00}, C_{10}, C_{11} , and C_{01} , for outcomes 1-4, respectively. It is 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 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 [16],

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

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 .

Because 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 express the risk in terms of the decision regions and probabilities as,

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

To find the decision rule, the decision regions are determined such that the risk in (3.2) is minimized. Because each element of \mathbf{y} 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 [16]

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

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

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

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

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

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 . 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 the points are assigned to A_1). We can, therefore, define the decision region in the observation space by

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

which can be rewritten as

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

The quantity on the left is called the *likelihood ratio* and will be denoted by

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

The relationship on the right is the threshold of the test and will be denoted by η .

Thus, Bayes criterion leads to a likelihood ratio test,

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

One of the methods for hypothesis testing is based on the Neyman-Pearson criterion [16]. In the Neyman-Pearson detection scheme the hypothesis test is formulated as a constrained optimization problem. In this optimization problem the false alarm probability is constrained and the probability of detection is maximized. The optimization problem yields a likelihood ratio test and thresholding conditions. The

Neyman-Pearson criterion [16], [35] generates a test to maximize P_d (probability of detection) while making P_{fa} (probability of false alarm) 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 [16].

We applied a block-based likelihood ratio test using the standard Gauss-Gauss detector [14] which is used to determine if a block belongs to a potential mine-like object or just background. The detection problem is viewed in terms of the signal plus noise model [14], the decision between two hypotheses is now either background (noise) only (H_0) or target (signal) plus background (H_1). Assuming that observation block of size $n \times n$ shaped column-wise into a vector $\mathbf{y} \in \mathbb{R}^N$ ($N = n^2$) is Gaussian distributed with zero mean and covariance matrix R . In the Gauss-Gauss detector, we 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 target (signal) alone. It is assumed that noise and target are uncorrelated. The conditional probability density function, $p_{\mathbf{Y}|H_i}(\mathbf{y}|H_i)$, for a given hypothesis H_i , $i \in [0, 1]$ and measurement vector \mathbf{y} is given by

$$p_{\mathbf{Y}|H_i}(\mathbf{y}|H_i) = (2\pi)^{-\frac{N^2}{2}} |R_i|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^H R_i^{-1} \mathbf{y}} \quad (3.10)$$

Now, using the likelihood ratio in (3.8) and taking the natural log, the log-likelihood of \mathbf{y} becomes [14]:

$$\begin{aligned} l(\mathbf{y}) &= \ln \left(\frac{(2\pi)^{-\frac{N^2}{2}} |R_1|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^H R_1^{-1} \mathbf{y}}}{(2\pi)^{-\frac{N^2}{2}} |R_0|^{-\frac{1}{2}} e^{-\frac{1}{2}\mathbf{y}^H R_0^{-1} \mathbf{y}}} \right) \\ &= \ln \left(\frac{|R_1|^{-\frac{1}{2}}}{|R_0|^{-\frac{1}{2}}} e^{\frac{1}{2}\mathbf{y}^H (R_0^{-1} - R_1^{-1}) \mathbf{y}} \right) \\ &= \frac{1}{2} \ln |R_1| - \frac{1}{2} \ln |R_0| + \frac{1}{2} \mathbf{y}^H (R_0^{-1} - R_1^{-1}) \mathbf{y} \end{aligned} \quad (3.11)$$

Disregarding the constants that are not observation dependent, the likelihood-ratio for the Gauss-Gauss detector [14] becomes

$$l(\mathbf{y}) = \mathbf{y}^H (R_0^{-1} - R_1^{-1}) \mathbf{y} = \mathbf{y}^H Q \mathbf{y}. \quad (3.12)$$

where $Q = R_0^{-1} - R_1^{-1}$.

Using this log-likelihood, the test in (3.9) is implemented for each block to determine if the block belongs to a mine-like object. Through our research we have found that the full-rank block-based method just described works well for detecting FOI, but yields incomplete ROI silhouettes making it difficult to classify detected ROI's.

Next, we will describe a process called rank-reduction [14] which maximizes the separation between targets and non-targets.

3.2.2 Rank Reduction

Let us start with (3.12) and rewrite matrix Q as:

$$Q = R_0^{-T/2}(I - S^{-1})R_0^{-1/2}. \quad (3.13)$$

where $R_0 = R_0^{1/2}R_0^{T/2}$ and $S = R_0^{-1/2}R_1R_0^{-T/2}$ is the “signal-to-noise ratio” matrix [14]. Under this transformation we can write the log likelihood ratio in (3.13) in terms of the “signal-to-noise ratio” matrix S as

$$l(\mathbf{z}) = \mathbf{z}^T(I - S^{-1})\mathbf{z} \quad (3.14)$$

where $\mathbf{z} = R_0^{-1/2}\mathbf{y}$ is also Gaussian distributed with zero mean and covariance matrix $R = I$ under H_0 and $R = S$ under H_1 i.e.,

$$\begin{aligned} E_{H_0}[\mathbf{z}\mathbf{z}^T] &= I \\ E_{H_1}[\mathbf{z}\mathbf{z}^T] &= S \end{aligned} \quad (3.15)$$

The J-divergence [14] which is a measure of the detectability (or separation) between the two hypotheses is written as

$$\begin{aligned} J &= E_{H_1}[l(\mathbf{y})] - E_{H_0}[l(\mathbf{y})] \\ &= \text{tr}(I - S^{-1})(E_{H_1}[\mathbf{z}\mathbf{z}^T] - E_{H_0}[\mathbf{z}\mathbf{z}^T]) \\ &= \text{tr}(S + S^{-1} - 2I) \end{aligned} \quad (3.16)$$

In order to maximize the J-divergence between H_0 and H_1 we look at the orthogonal decomposition of the S matrix:

$$S = R_0^{-1/2} R_1 R_0^{-T/2} = U \Lambda U^T. \quad (3.17)$$

where Λ is a diagonal matrix with diagonal elements λ_i and U is an orthogonal matrix satisfying $UU^T = I$. In this form, the log likelihood ratio becomes

$$l(\mathbf{y}) = \mathbf{z}^T U (I - \Lambda^{-1}) U^T \mathbf{z} \quad (3.18)$$

and the J-divergence between the two hypothesis becomes

$$\begin{aligned} J &= \text{tr}(\Lambda + \Lambda^{-1} - 2I) \\ &= \sum_{i=1}^N (\lambda_i + \lambda_i^{-1} - 2) \end{aligned} \quad (3.19)$$

As can be seen both the log-likelihood in (3.18) and J-divergence in (3.19) are now expressed in terms of the eigenvalues and eigenvectors of the SNR matrix S . Also in (3.19) it is obvious that it is the sum of $(\lambda_i + \lambda_i^{-1} - 2)$ that determines the contribution to the J-divergence. It can be shown (see Remark 1 below) that the term $(\lambda_i + \lambda_i^{-1} - 2)$ determines the best per-mode SNR contribution to the J-divergence. This means that eigenvalues that are either much larger than unity or much less than unity should be retained for best case rank reduction, and improvement of SNR.

Equation (3.18) can be written in the reduced-rank form by using Λ_r and I_r instead of Λ and I , where Λ_r and I_r contain only r non-zero elements along the diagonals (the r nonzero entries occur at arbitrary locations) as defined in [14]. Matrices I_r and

Λ_r^{-1} are

$$I_r = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 0 \end{bmatrix}$$

$$\Lambda_r^{-1} = \begin{bmatrix} \lambda_1^{-1} & & & & & \\ & \lambda_2^{-1} & & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & \lambda_r^{-1} & \\ & & & & & 0 \end{bmatrix}$$

The reduced rank log likelihood ratio and J-divergence then become

$$l(\mathbf{y}) = \mathbf{z}^T U (I_r - \Lambda_r^{-1}) U^T \mathbf{z} \quad (3.20)$$

and

$$\begin{aligned} J_r &= \text{tr}(\Lambda_r + \Lambda_r^{-1} - 2I) \\ &= \sum_{n=1}^r (\lambda_n + \lambda_n^{-1} - 2) \end{aligned} \quad (3.21)$$

The reduced rank log-likelihood ratio in (3.20) is used in our reduced-rank block-based detector which maximizes the J-divergence, between the two hypotheses H_0 and H_1 . A procedure is suggested in [14] for choosing only a subset of r eigenvalues of S to maximize the J-divergence in (3.21). We have found experimentally that reducing the detector rank to $r = 1$ always yields the highest separation between the two hypotheses. This process in essence reduces the effects of background noise in the detection process.

Remark 1 If $R_0 = \sigma^2 \mathbf{I}$ and $R_s = \text{diag}[\sigma_{s_1}^2 \cdots \sigma_{s_N}^2]$, then

$$\lambda_i = \frac{\sigma_{s_i}^2 + \sigma_n^2}{\sigma_n^2} \quad (3.22)$$

and thus

$$(\lambda_i + \lambda_i^{-1} - 2) = \frac{\sigma_{s_i}^4}{\sigma_n^2(\sigma_{s_i}^2 + \sigma_n^2)} = \frac{SNR_i^2}{SNR_i + 1} \approx SNR_i \quad (3.23)$$

where $SNR_i = \frac{\sigma_{s_i}^2}{\sigma_n^2}$. That is, each term in (3.21) corresponds to the “per-mode” SNR.

3.2.3 Implementation of Block-based Detection

Each frame in a data run is partitioned into blocks of size 4×4 ($n = 4$). Each block is then rearranged into a N-dimensional column vector ($N = n^2 = 16$) for computing the log-likelihood ratio.

An exaggerated example of the blocking is shown in Figure 3.1. Each block is column-wise rearranged into a vector in order to compute the log-likelihood ratio.

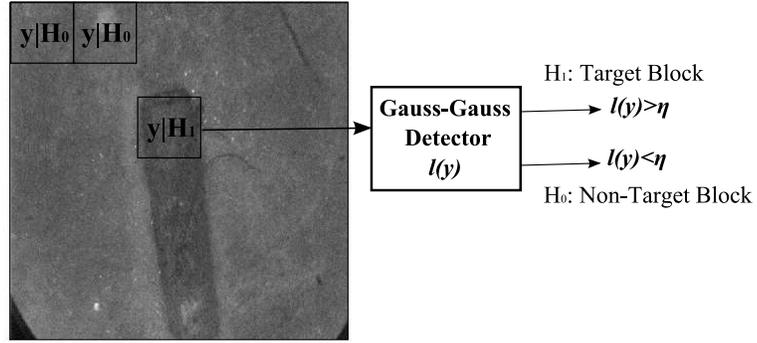


Figure 3.1: Block-based detection process.

The Gauss-Gauss detection is then performed on each block, and a likelihood value is computed for each block which generates a ‘likelihood map’. In this likelihood map each pixel represents the value of the log-likelihood ratio of the corresponding block in the original EO image. The likelihood maps are then used to determine both FOI, and

ROI in a data set based on thresholding the log-likelihood ratio. This log-likelihood ratio thresholding is based on the training data.

Size thresholding is also imposed on the number of detected blocks needed in order to declare a frame as a FOI. If only isolated blocks are detected in a frame it does not warrant calling the frame a FOI. Thus, in our implementation at least 180 connected blocks must be detected in order for the frame to be flagged as a FOI. Also an upper size threshold of 2500 connected blocks is imposed. If too many connected blocks are detected then the frame is assumed to contain only background anomalies.

These size constraints pose another challenge to the FOI detector. More specifically, if the vehicle carrying the sensor is high above the targets, the targets may appear small, and may be missed due to the lower size threshold, conversely if the vehicle is directly over the target then the target may appear too large. The above size constraints were determined experimentally using the different mine-like objects in the database. The overall process is described in the following steps:

1. Extract target, and background blocks from the training set from both mine-like objects and background anomalies frames in order to determine a threshold value for separating the likelihood value of the target and non-target blocks.
2. Compute the likelihood ratio for each block under the Gauss-Gauss formulation using (3.20).
3. Threshold the likelihood ratio for each block. If the block's likelihood ratio falls above the threshold then designate the particular block as 'target'. If the block's likelihood ratio is below the threshold then designate the block as background.
4. The number of connected blocks is determined using MATLAB `regionprops` function. If a particular number of connected blocks are designated as a target blocks, then flag the frame under consideration as a FOI.

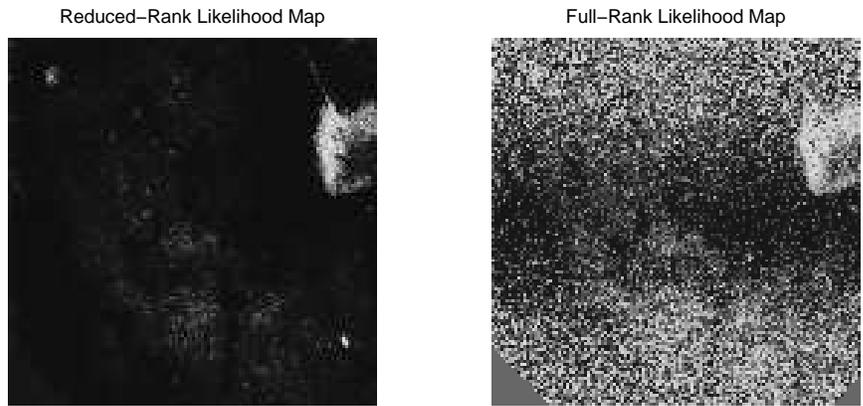
5. The ROI is automatically determined and segmented directly as a result of this process by way of the detected blocks in the FOI. This is because the collection of connected block form the silhouette of the object.

3.2.4 Comparison Between Reduced-Rank and Full-Rank Detectors

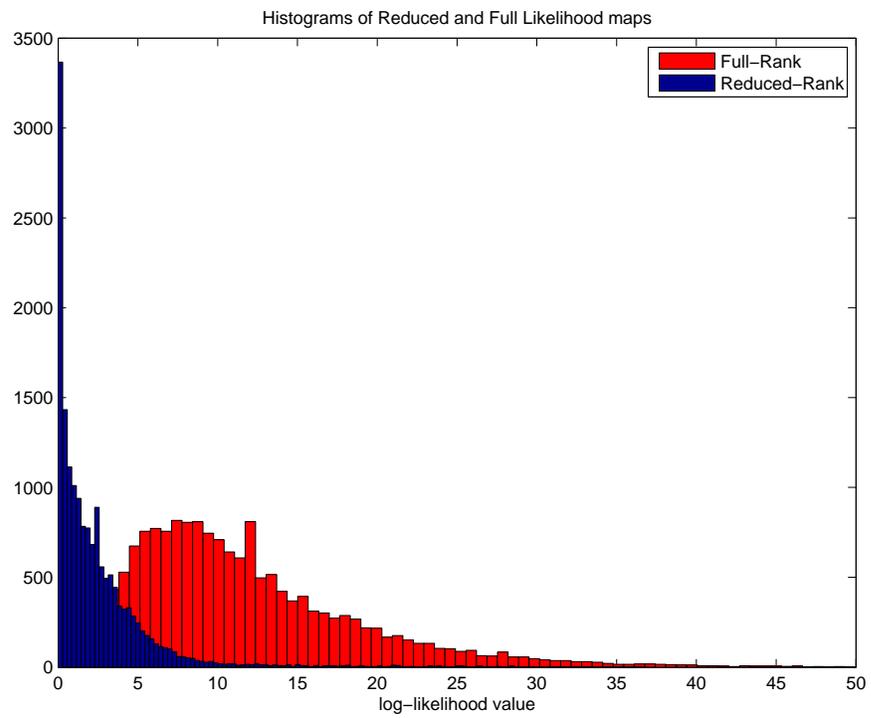
As mentioned before, for the reduced rank detector $r = 1$ gave the best separation between log-likelihood ratio values for H_0 and H_1 . An example of this separation is shown in Figure 3.2(a), which shows the log-likelihood maps for these detectors. A comparison between the log-likelihood ratio values is shown in the histograms in Figure 3.2(b) for the two cases. It can be seen from Figures 3.2(a) and 3.2(b) that for reduced rank detector, values of the likelihood have been pushed towards lower values. This has in turn suppressed much of the noise from the full rank implementation, and hence improved the SNR. Therefore, we choose to work with the reduced rank version of the log likelihood ratio test in our overall system.

3.2.5 Detector Design

In order to use the proposed block-based likelihood detector first a ‘training set’ must be selected. The selection of a set of blocks from some mine-like objects and background anomalies is required in order to compute the covariance matrices associated with H_0 and H_1 . This process is subjective in that the blocks used for the training must be hand picked from frames which are believed to represent a wide range of target and background scenarios. For this purpose we have designed a software application GUI (see Appendix A for details) which aids in the selection of training blocks for the detector, as well as building a feature set for training the neural network classifier. Since a limited number of frames containing targets are available in this database, blocks from two frames in a single data run (SAM001_004) containing a target were used. In order to form the training set, regions of target and background were cropped from the frames shown in Figure 3.3 using the developed



(a) Comparison between full rank and reduced rank log likelihood ratio maps. Both maps are plotted on the same scale from 0 (black) to 50 (white)



(b) Comparison between histograms of the full rank and reduced rank log likelihood ratio values.

Figure 3.2: Comparison between full rank and reduced rank detectors.

GUI software application. That is, we construct a mine-like object training matrix $\mathbf{T} = [\underline{T}_1, \underline{T}_2, \dots, \underline{T}_K]$ where the subscripts denote the block index and are obtained from several different blocks over mine-like objects. The subscript K is the total number of *training* blocks used, which should be large enough to contain the variety of mine-like object signatures that are typically encountered. We also construct a background clutter training matrix $\mathbf{N} = [\underline{N}_1, \underline{N}_2, \dots, \underline{N}_K]$ which are obtained from several different blocks containing only background anomalies. For both mine-like objects, and background cases $K = 1465$ blocks which should be large enough to contain the variety of mine-like object and background scenarios that are typically encountered.

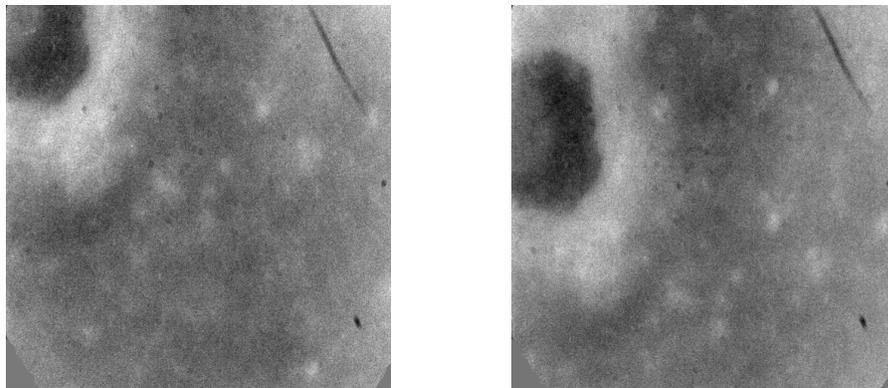


Figure 3.3: FOI from data run SAM001_004 used to train the Gauss-Gauss detector. Regions of blocks were selected over the target and over background.

Blocks were cropped from two frames in the data run SAM001_004 for each mine-like objects and background anomalies to compute covariance matrices R_1 and R_0 , respectively. The 1465 training blocks for target contained mostly pixels over the target, and target edges, while the 1465 randomly chosen blocks for background contained only pixels belonging to background anomalies. After training the detector was evaluated on the data set in Table 2.1 in order to assess the system’s performance for FOI detection.

A fixed threshold is chosen based on several frames in a ‘detector validation run’ SAM002_008. The detector has been evaluated on several frames containing mine-like objects in run SAM002_008 in order to determine a suitable threshold to successfully differentiate between background and target blocks. Using the distributions of log-likelihood values for target and background blocks in SAM002_008 we have experimentally chosen the threshold to be 5. It turns out that this threshold is adequate for detecting mine-like objects in the testing set considered in this study.

3.3 ROI Segmentation

After each image has been passed through the detector and the likelihood map has been generated the likelihood values are thresholded. This thresholding process is used to segment all detected objects in a given frame. Each block is compared to a threshold η as in Figure 3.1. If a block’s likelihood value lies above the determined threshold the block is designated as a ‘target block’. If a block’s likelihood value lies below the threshold η the block is designated as a ‘background block’. The detector is somewhat robust to rotation of ROI due to the fact that the interior blocks of the ROI are mainly detected. However, edge blocks may be affected by rotation depending on the degree. Thus, it is possible that detection of partial targets could be affected by rotation depending on the severity of occlusion in the frame. Missed detections may occur if there are not enough interior blocks of a target in a given FOI to flag a

detection.

After the thresholding process is completed a binary image remains containing collections of target blocks which constitute 'target objects'. Each detected target object in a given frame is now compared to an upper and lower 'object size threshold'. The object size threshold imposes upper and lower size constraints on the number of connected blocks on the detected target objects. Figure 3.4 shows a block diagram outlining the overall detection and ROI segmentation process in this block-based detection scheme. Figure 3.4 shows that we start with the original image which is passed through the detector to generate the likelihood map. Once the likelihood map has been generated we threshold the pixel intensity values of the likelihood map to generate the binary image. Once the binary image has been generated we then impose the object size threshold to remove objects that are assumed too large or too small to be a potential mine-like object. If objects are found to be within the size constraints then the frame is flagged as FOI. Lastly, the remaining binary silhouettes are collected to be further processed by the classifier to discriminate mine-like object from background clutter using shape-based features described in Chapter 5. If a single object is detected in a given frame, the frame is flagged to be an FOI, otherwise if no objects are detected in the frame then the frame is discarded. Each detected object in a given FOI is flagged to be an ROI. Ultimately the detector outputs binary silhouettes of the detected objects as well as several measures associated with the detected objects, which are summarized in Table 3.1.

The detection results are output to a new folder based on the date and time the detector is run. The detection measures are output to a text file which contain the information in Table 3.1 formatted in rows. The detector also saves figures of the detected FOI with ROI's bounding boxes, and binary silhouettes of the detected objects.

Several important measures associated with each detection are given in Table 3.1.

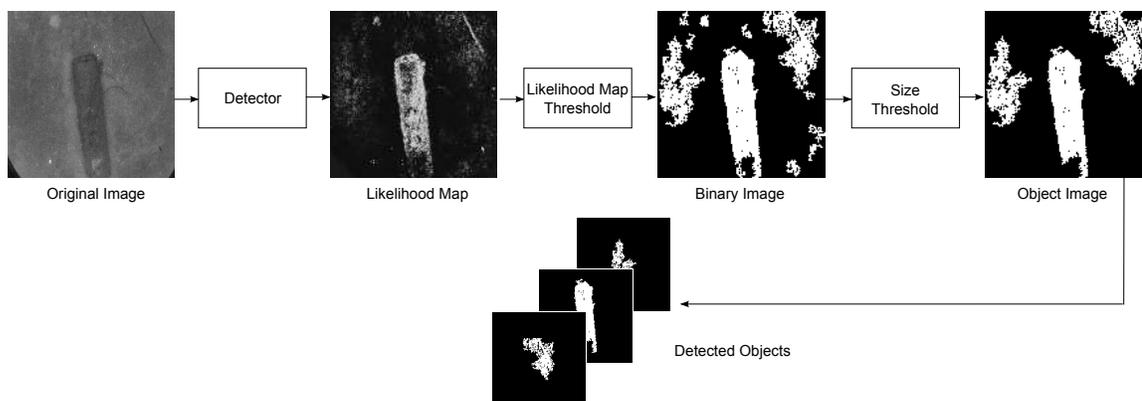


Figure 3.4: Detection and ROI segmentation process.

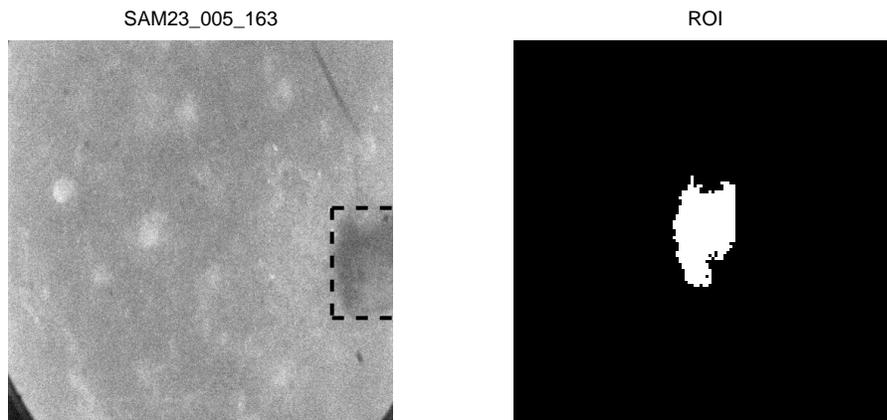
First the **Frame** is given to indicate which frame the detection has occurred in. Next, **ObjectNumber** is given to indicate whether or not several objects have been detected in a single FOI. The **Area** measure is used to report the number of connected blocks contained in each detected object, which gives an indication of the size of the detected object. **CentroidX** and **CentroidY** report the x and y center position of each detected object (or ROI) within the FOI. **MajorAxisLength** and **MinorAxisLength** specifies the length (in pixels) of the major, and minor axis of the ellipse that has the same normalized second order central moments as the ROI, respectively. The **Eccentricity** measure specifies the eccentricity of an ellipse that has the same second-moments as the ROI, while **Orientation** reports the angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of an ellipse that has the same second-moments as the ROI. **Solidity** indicates the proportion of the pixels in the convex hull that are also in detected ROI, which is computed as $\text{Area}/\text{ConvexArea}$. **Xpos** and **Xwid** report the coordinate (pixel column) and horizontal width (pixels) of the bounding box of the detected ROI, respectively, while **Ypos** and **Ywid** report the coordinate (pixel row) and vertical width (pixels) of the bounding box of the detected ROI, respectively.

Several detection results (output figures) are shown for targets in Figures 3.5(a)-3.5(c) and background anomalies in the Figures 3.6(a)-3.6(c). In each of these figures

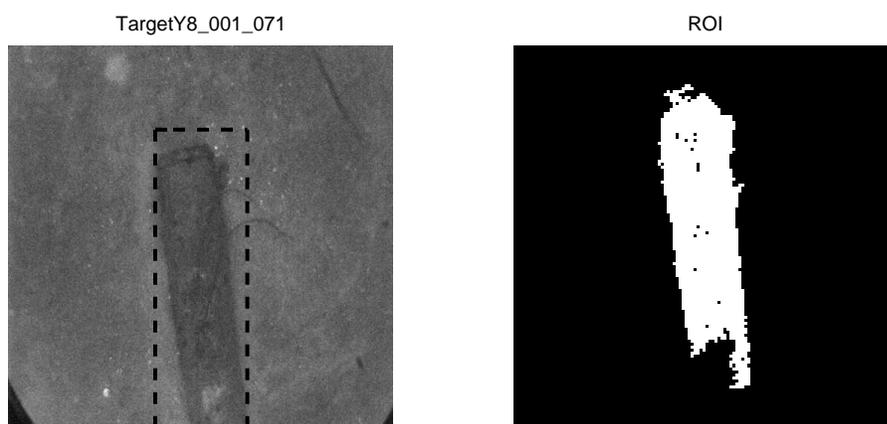
Table 3.1: Detector output Measures

Measure	Description
Frame	Frame number associated with the input file (given by the last three numeric digits in the *.tif file name).
ObjectNumber	The ROI number of the detected object in the frame.
Area	The number of detected blocks in the ROI.
CentroidX	The horizontal coordinate (pixel column) location of the ROI center of mass.
CentroidY	The vertical coordinate (pixel row) of the ROI center of mass.
MajorAxisLength	Specifies the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the ROI.
MinorAxisLength	Specifies the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the ROI.
Eccentricity	Specifies the eccentricity of the ellipse that has the same second-moments as the ROI
Orientation	The angle (in degrees ranging from -90 to 90 degrees) between the x-axis and the major axis of the ellipse that has the same second-moments as the ROI.
Solidity	The proportion of the pixels in the convex hull that are also in the ROI. Computed as Area/ConvexArea.
NetScore	Score generated by the Neural network shape based classification.
Xpos	The horizontal coordinate (pixel column) of the bounding box of the ROI.
Ypos	The vertical coordinate (pixel row) of the bounding box of the ROI.
Xwid	The horizontal width (pixels) of the bounding box of the ROI.
Ywid	The vertical width (pixels) of the bounding box of the ROI.

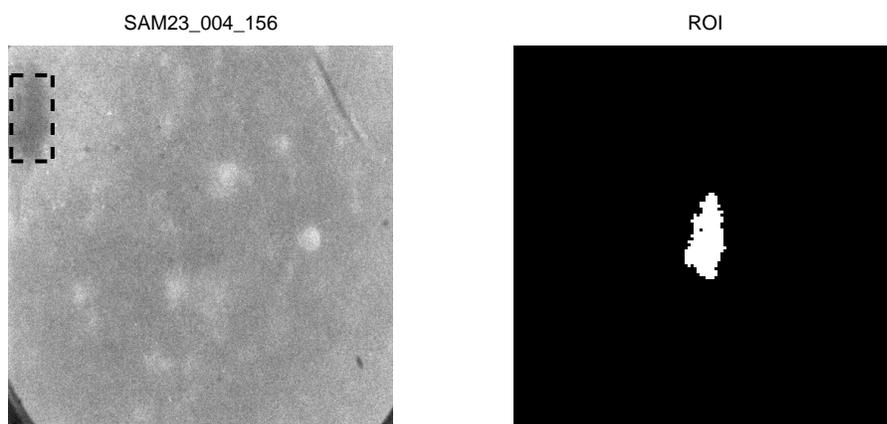
the original image is shown on the left and segmented ROI from the likelihood map is shown on the right. The bounding box of the ROI is superimposed on the original input frame. It can be seen that the detector generates well-defined silhouettes for the target cases. It is also important to note that for the background anomalies detected the silhouettes are irregularly shaped with more holes. Table 3.2 shows a summary of FOI detections at the point of 100% detection of targets. Overall, an FOI detection rate of 100% has been achieved at the cost of a FOI false alarm rate of 36.9% when considering data in Table 2.1. This means that 486/1317 of the frames in the data set pass the detector, which will be subsequently applied to the classifier to classify the detected ROI and further reduce the false alarm rate. We can see the runs have varying degrees of false alarms due to varying ocean bottom conditions in each run. Typically, more false alarms are obtained when the ocean bottom has more dense clutter.



(a) Partial Target SAM23.005 Frame 163

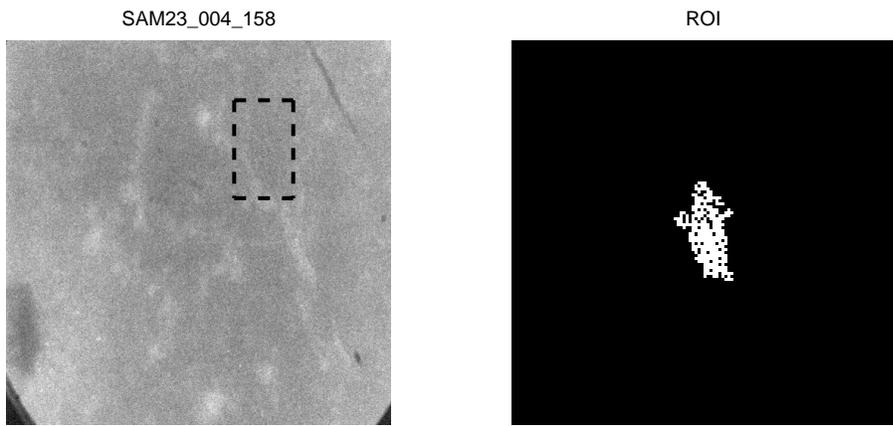


(b) Target TargetY8.001 Frame 071

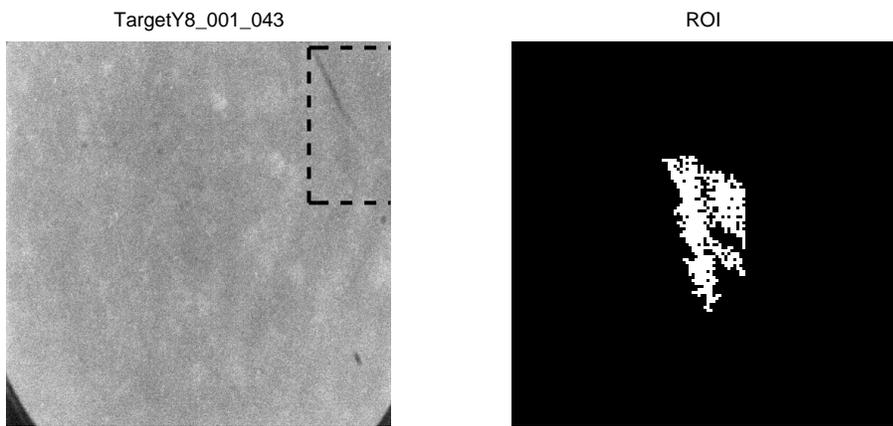


(c) Partial Target SAM23.004 Frame 156

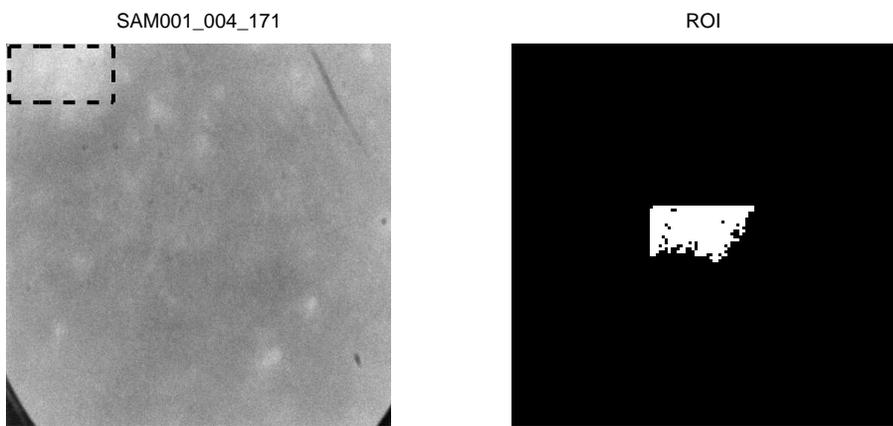
Figure 3.5: Various detector outputs for different target frames.



(a) False Alarm (with target) SAM23.004 Frame 158



(b) False Alarm TargetY8.001 Frame 043



(c) False Alarm SAM001.004 Frame 171

Figure 3.6: Various detector outputs for different detected background anomalies.

Table 3.2: Detection Results for Reduced-Rank Gauss-Gauss

Run	False Detections	True Detections/FOI Total
SAM001_003	2 / 42	(0 / 0)
SAM004_001	7 / 35	(0 / 0)
SAM22_011	12 / 35	(0 / 0)
SAM23_003	142 / 293	(3 / 3)
SAM23_004	92 / 287	(3 / 3)
SAM23_005	129 / 293	(4 / 4)
TargetY8_001	47 / 136	(3 / 3)
TargetY8_003	11 / 29	(0 / 0)
TargetY8_004	16 / 32	(0 / 0)
TargetY8_006	28 / 135	(3 / 3)
Totals	486/1317	16/16
Percentage	36.9%	100%

3.4 Conclusion

In this chapter we described and analyzed a full-rank and a reduced-rank block-based detector for detecting underwater mine-like objects. Using block sizes of 4x4 the log-likelihood ratio has been evaluated on the entire EO-CCD database. First it was found that a detector that used 4x4 blocks performed much better than that used 8x8 sized blocks. Second it has been observed that detected ROI's of targets are more regularly shaped and are more solid than those of the background anomalies. Additionally the detection algorithm does not require any preprocessing, i.e., it is directly applied to the data acquired from the EO-CCD sensor. We have found that a challenge with this new data set is the fact that FOI must be identified before segmentation and ROI feature extraction. Also, detection of partial targets and classification based on partial silhouettes pose many challenges. A promising benefit of the proposed block-based detector lies in the fact that detection of FOI, and segmentation can be achieved in a single step. This automatic FOI detection coupled with automatic ROI segmentation and object size thresholding reduces the number of objects from which features need to be extracted, and hence reducing the load on the classifier. Automatic FOI detection and ROI segmentation are desirable benefits of this detection scheme.

Overall, robust object silhouette definition has been achieved using the reduced-rank Gauss-Gauss detector. Silhouette definition is the key to successful object classification. Using methods described in the next chapter it is found that the matched subspace detection method provides further improvement in clutter suppression and hence further reduces the false alarm rate when compared to that of the reduced-rank Gauss-Gauss detector.

CHAPTER 4

MATCHED SUBSPACE DETECTION

4.1 Introduction

In Chapter 3 we showed that in order to determine FOI in a data run, each frame is partitioned into small blocks (e.g. of size 4×4) and the problem is viewed as block-based binary hypothesis test under the Neyman-Pearson criterion [14, 16]. We showed how block-based detection methods provide automatic ROI segmentation of the detected objects. Reduced-rank version of the Gauss-Gauss detector [14] was also presented, implemented and benchmarked for the data set in Table 2.1. This reduced-rank detector improves the detection performance by suppressing the noise/clutter. In this chapter we present methods for further suppressing background clutter from EO-CCD imagery to facilitate more accurate object detection, at lower computational cost comparing to the reduced-rank Gauss-Gauss detector. Mitigating the effects of background anomalies allows for better extraction of target ROI while reducing false detections hence reducing the number of objects that need to be classified. This reduction in false alarm lowers the computational load on the classifier which is important for real time applications.

We have investigated those methods for background removal that are based on the framework in [15], [17] which outlines various forms of the matched subspace detection under different assumptions for the structure of the signal, clutter, and noise. In particular, a method for separating the background clutter by projecting the data onto the signal subspace will be introduced. In this chapter we first review the theory of the matched subspace detection in [15], [17] for the signal plus noise structure. Next, we give a comparison between the reduced-rank Gauss-Gauss detector, and

matched subspace detector on the data contained in Table 2.1.

This chapter is organized as follows: In Section 4.2 we present the theory of the matched subspace detector. In Section 4.3 we present a comparison between the reduce-rank Gauss-Gauss detector and show improved clutter suppression can be achieved using the matched subspace method, which uses the same block-based framework for FOI, and ROI detection. Section 4.4 gives a summary of this chapter and concluding remarks on the detector.

4.2 Matched Subspace Detection

As previously described in Chapter 3 we use the likelihood ratio test (LRT) to determine an unknown hypothesis (H_i) from a known measurement. This method for hypothesis testing is based on the Neyman-Pearson criterion which yields the LRT and optimum thresholding conditions [16]. The method of matched subspace detection utilizes the generalized likelihood ratio test (GLRT) in which the unknown parameter we wish to estimate is replaced with its maximum likelihood (ML) estimate. In this case, the GLRT is a maximal invariant statistic which means that the GLRT is the uniformly most powerful invariant detector [15], [17]. In [15], the problem is to determine whether observation \mathbf{y} is associated with signal plus noise plus interference, or with noise plus interference conditions. Several cases of detecting subspace signals in the presence of subspace interference and noise are considered in [15]. While reference [17] considers the problem of detecting subspace signals in noise (without interference).

In the signal plus noise (without structured interference) matched subspace framework we regard the signal (mine-like object) \mathbf{x} to be represented as a linear combination of modes or basis vectors [17]. The motivation for using this signal model in our problem is based on the fact that each measurement block \mathbf{y} in a given EO-CCD image will contain either a target's signature plus background noise $\mathbf{x} + \mathbf{n}$ under H_1

or background noise alone \mathbf{n} under H_0 . We make the assumption that the mine-like objects we wish to detect can be characterized using a small number of basis vectors e.g., blocks corresponding to different edge orientations and texture content. There are several ways to formulate the matched subspace detection problem. We use the framework presented in [17] due to the fact that we lack structured interference. As before, the binary hypothesis test is written as:

$$\begin{aligned} H_1 &: \mathbf{y} = \mathbf{x} + \mathbf{n} \\ H_0 &: \mathbf{y} = \mathbf{n} \end{aligned}$$

As with the Gauss-Gauss detectors in Chapter 3 each EO-CCD image is first divided into blocks of size $n \times n$ and arranged (column-wise) into a vector of $N \times 1$ ($N = n^2$) denoted by \mathbf{y} . The background removal method requires knowledge of the signal subspace $\langle \mathbf{H} \rangle$ (e.g. target blocks \mathbf{x} spanning $\langle \mathbf{H} \rangle$), which is of dimension $M_{\mathbf{H}} < N$. This requires finding $M_{\mathbf{H}}$ suitable basis vectors that properly span $\langle \mathbf{H} \rangle$. $M_{\mathbf{H}} < N$ basis vectors can be obtained by applying singular value decomposition (SVD) [36] to a data matrix that has a large number of training block vectors as its columns. That is, we first construct a mine-like object training matrix $\mathbf{T} = [\underline{\mathbf{T}}_1, \underline{\mathbf{T}}_2, \dots, \underline{\mathbf{T}}_K]$ where each column corresponds to a different block which contain mine-like objects. The subscript K is the total number of *training* blocks should be large enough to contain the variety of mine-like object signatures that are typically encountered. Note this matrix \mathbf{T} is the same training matrix used to form the sample covariance matrix for targets R_1 in Chapter 3. Specifically, $K = 1465$ blocks were cropped from two frames in the data run SAM001_004 which are typically encountered mine-like objects.

Now, we assume the signal \mathbf{x} obeys a linear subspace model.

$$\mathbf{x} = \sum_{n=1}^p \mathbf{h}_n \theta_n = \mathbf{H}\boldsymbol{\theta} \quad (4.1)$$

$$\mathbf{H} \in \mathbb{R}^{N \times p}, \boldsymbol{\theta} \in \mathbb{R}^p$$

Here \mathbf{H} is a known $N \times p$ matrix with columns \mathbf{h}_n and θ is a $p \times 1$ vectors with elements θ_n . The matrix \mathbf{H} is composed of $p \leq K$ basis vectors which represent the signal subspace $\langle \mathbf{H} \rangle$. One way to define matrix \mathbf{H} is to apply SVD on the data matrix \mathbf{T} to give a representative set of basis vectors for mine-like objects, i.e.,

$$\mathbf{T} = \mathbf{U}\Sigma\mathbf{V}^T \quad (4.2)$$

where the columns of \mathbf{U} contain the input basis vectors for \mathbf{T} , the columns of \mathbf{V} contain the output basis vectors for Y , and Σ is a diagonal matrix of singular values. The first $M_{\mathbf{H}}$ columns of \mathbf{U} corresponding to the $M_{\mathbf{H}}$ largest singular values of \mathbf{T} are used to form \mathbf{H} (i.e. basis vectors that span $\langle \mathbf{H} \rangle$) since this allows reconstruction of the data vectors in \mathbf{T} optimally in the least squares sense. In this experiment we select six columns of \mathbf{T} corresponding to the largest singular value of Σ . To formulate the matched subspace detector we first write our hypothesis H_0 and H_1 in terms of the corresponding noise models, i.e.,

$$\begin{aligned} H_1 & : \mathbf{n}_1 = \mathbf{y} - \mathbf{H}\theta \\ H_0 & : \mathbf{n}_0 = \mathbf{y} \end{aligned} \quad (4.3)$$

where \mathbf{n}_1 and \mathbf{n}_0 are noise under hypothesis H_0 and H_1 respectively. Here we make the assumption that \mathbf{n}_1 and \mathbf{n}_0 are Gaussian distributed random variables with diagonal covariance structures with variance σ^2 . To justify this assumption the sample covariance matrices for data matrices \mathbf{N} and \mathbf{T} are generated and shown in Figure 4.1. As can be seen, for data matrix \mathbf{N} the diagonal elements are much larger than the off-diagonal ones. This assumption implies that different pixels within a block are almost uncorrelated with one another. While this may be a reasonable assumption for \mathbf{n}_0 , clearly it is not accurate for \mathbf{n}_1 based on the covariance matrix for \mathbf{H}_1 having large off-diagonal elements as shown in Figure 4.1. Nonetheless, this simplifies the derivations significantly.

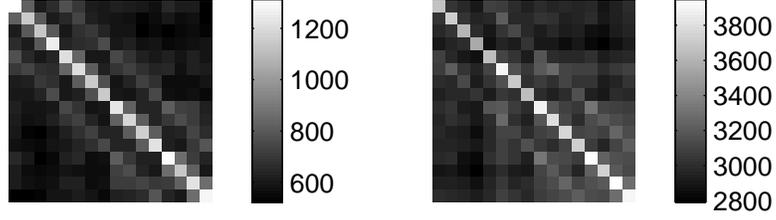


Figure 4.1: Covariance structure computed using \mathbf{N} (left) and \mathbf{T} (right) data matrices. These matrices show large values along the diagonal and lower values off the diagonal.

Assuming the noise is Gaussian distributed the probability density functions for \mathbf{n}_1 and \mathbf{n}_0 are given by

$$f_i(\mathbf{y}) = (2\pi\sigma^2)^{-N/2} \exp\left\{-\frac{1}{2\sigma^2} \|\mathbf{n}_i\|_2^2\right\}, i = 0, 1 \quad (4.4)$$

and the likelihood ratio function is defined as

$$l(\mathbf{y}) = \ln \frac{f_1(\mathbf{y})}{f_0(\mathbf{y})} = \|\mathbf{n}_0\|_2^2 - \|\mathbf{n}_1\|_2^2 \quad (4.5)$$

In order to formulate the GLRT we replace θ in (4.3) with it's ML estimate $\hat{\theta} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y}$ hence giving

$$\begin{aligned} l(\mathbf{y}) &= \|\hat{\mathbf{n}}_1\|_2^2 - \|\hat{\mathbf{n}}_2\|_2^2 \\ &= \mathbf{y}^T \mathbf{y} - (\mathbf{y} - \mathbf{P}_{\mathbf{H}} \mathbf{y})^T (\mathbf{y} - \mathbf{P}_{\mathbf{H}} \mathbf{y}) \end{aligned} \quad (4.6)$$

Computing the right hand side of 4.6 yields the following GLRT:

$$l(\mathbf{y}) = \mathbf{y}^T \mathbf{P}_{\mathbf{H}} \mathbf{y} \quad (4.7)$$

where $\mathbf{P}_{\mathbf{H}} = \mathbf{H}(\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$ is the projection matrix [36] onto the subspace $\langle \mathbf{H} \rangle$. The term $\mathbf{P}_{\mathbf{H}} \mathbf{y}$ in (4.7) is the component of our measurement \mathbf{y} that lies in the signal subspace $\langle \mathbf{H} \rangle$. Equation (4.7) gives the power of the measurement in the subspace $\langle \mathbf{H} \rangle$. In this experiment we have implemented matched subspace detector GLRT, and

evaluated it on the data set in Table 2.1. A comparison between the matched subspace detector and the Gauss-Gauss detectors in Chapter 3 is given in the following section.

4.3 Comparison between Matched Subspace & Gauss-Gauss Detection

In this section we present a comparison between the reduced-rank Gauss-Gauss detector of Chapter 3, and the matched subspace detector presented in the previous section. We have benchmarked both detectors on the dataset contained in Table 2.1. Here we present a comparison between the results of the three different steps (likelihood map, binary image, and object (ROI) image) of the reduced-rank Gauss-Gauss, and the matched subspace detector. Figures 4.2, 4.3, and 4.4 show the results of various detection stages for both reduced-rank Gauss-Gauss and matched subspace detectors. The figures shown from top to bottom are: the original image, likelihood map, binary image, and object (ROI) image for the three target run examples (SAM23_005, TargetY8_001, SAM23_004) presented in Section 3.3.

In each of these figures we see that the likelihood map has considerably better contrast for the case of the matched subspace detector. We can also see from these figures that the matched subspace likelihood map better separates the target from the background clutter. Figures 4.2, 4.3, and 4.4 also show the binary images for both reduce-rank Gauss-Gauss and matched subspace detectors obtained at a threshold corresponding to the 100% detection point on the ROC curve (Figure 4.5) of 100% true target detection for both detectors. We can see that the matched subspace detector provides improved clutter suppression in all three cases having lesser false alarms/detected clutter. The matched subspace method reduces clutter by measuring the energy of each block in the signal subspace, as opposed to the Reduced-Rank Gauss-Gauss detector which relies on the separation between second order statistics.

Tables 4.1 and 4.2 show the FOI detection results at the point of 100% detection

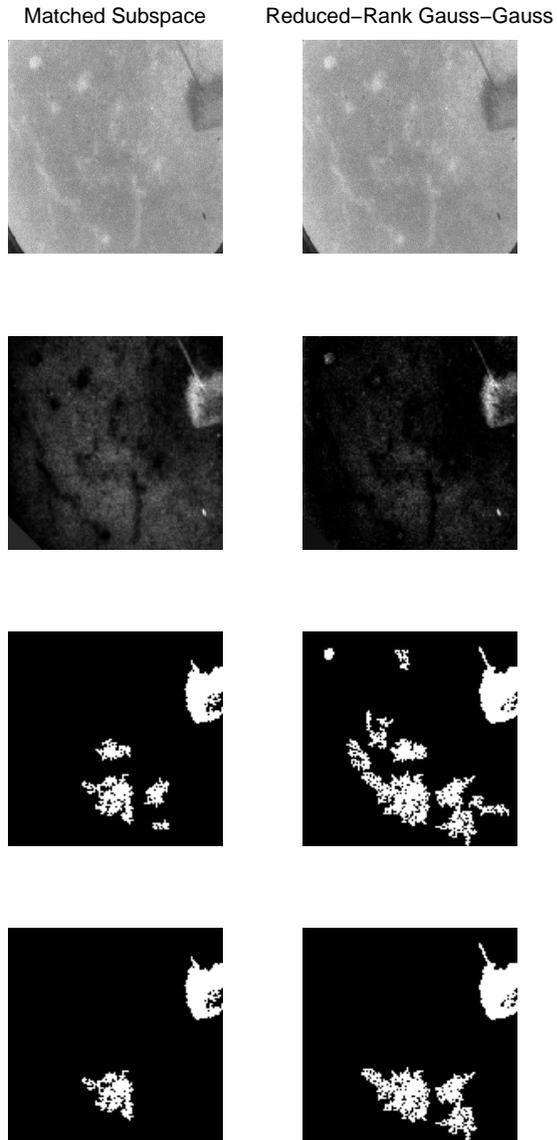


Figure 4.2: Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run SAM23_005 frame 162

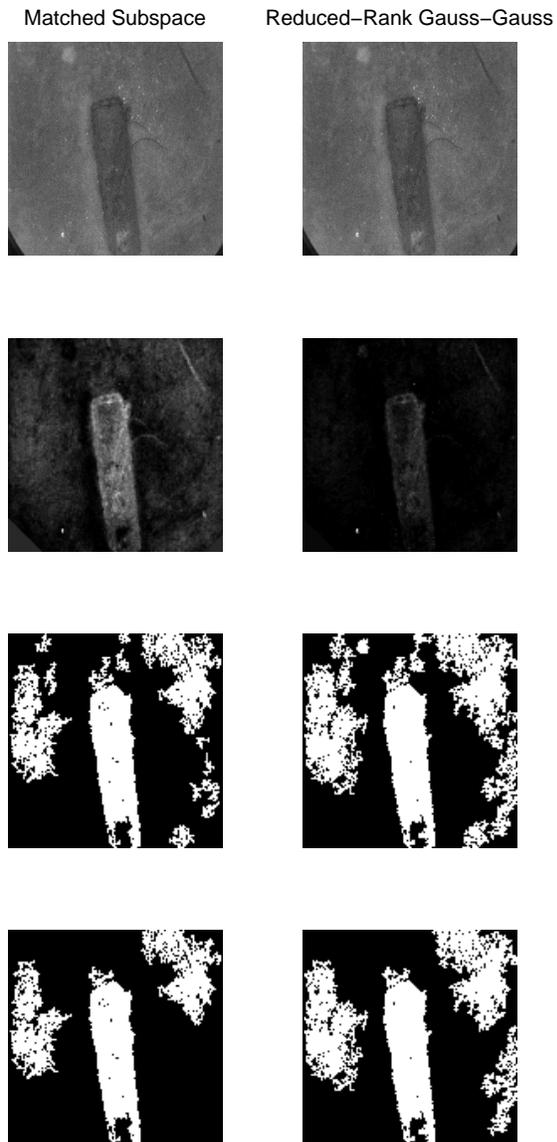


Figure 4.3: Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run TargetY8_001 frame 71

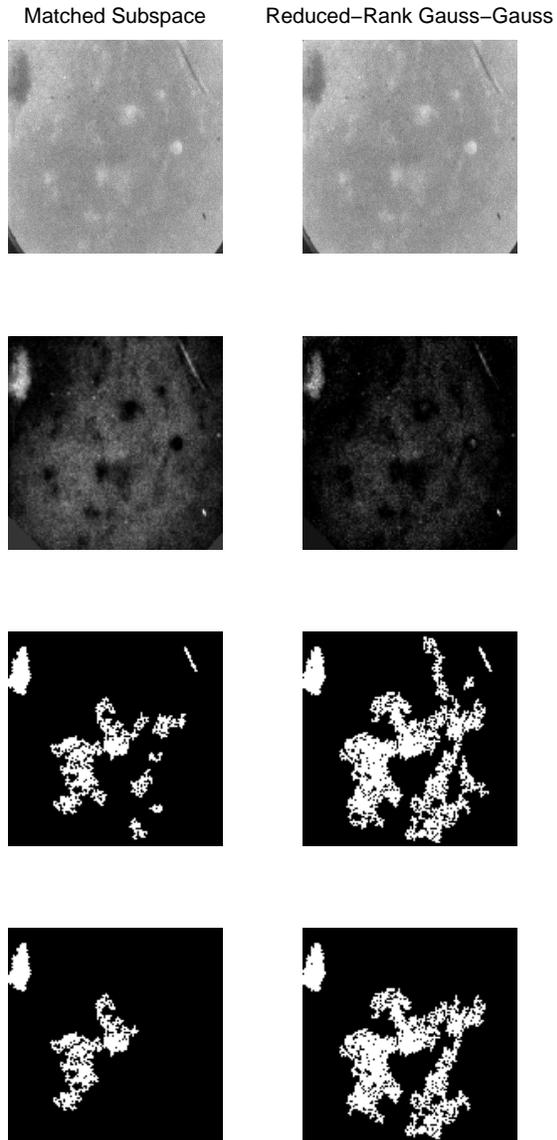


Figure 4.4: Comparison between matched subspace detector and reduce-rank Gauss-Gauss detector (from top to bottom) for the original image, likelihood map, binary image, and object image for target run SAM23_004 frame 156

for each detector. We can see from these tables that the matched subspace detector reduces the number of frames which need to be passed to the classifier by 16.7% from the reduced-rank Gauss-Gauss detector. This improvement is gained without increasing the computational time of the detector.

Table 4.1: Detection Results for Matched Subspace Detector

Run	False FOI Detections	True FOI Detections/FOI Total
SAM001_003	3 / 42	(0 / 0)
SAM004_001	0 / 35	(0 / 0)
SAM22_011	9 / 35	(0 / 0)
SAM23_003	60 / 293	(3 / 3)
SAM23_004	62 / 287	(3 / 3)
SAM23_005	63 / 293	(4 / 4)
TargetY8_001	31 / 136	(3 / 3)
TargetY8_003	9 / 29	(0 / 0)
TargetY8_004	1 / 32	(0 / 0)
TargetY8_006	28 / 135	(3 / 3)
Totals	266/1317	16/16
Percentage	20.2%	100%

Table 4.2: Detection Results for Reduced-Rank Gauss-Gauss Detector

Run	False FOI Detections	True FOI Detections/FOI Total
SAM001_003	2 / 42	(0 / 0)
SAM004_001	7 / 35	(0 / 0)
SAM22_011	12 / 35	(0 / 0)
SAM23_003	142 / 293	(3 / 3)
SAM23_004	92 / 287	(3 / 3)
SAM23_005	129 / 293	(4 / 4)
TargetY8_001	47 / 136	(3 / 3)
TargetY8_003	11 / 29	(0 / 0)
TargetY8_004	16 / 32	(0 / 0)
TargetY8_006	28 / 135	(3 / 3)
Totals	486/1317	16/16
Percentage	36.9%	100%

Figure 4.5 shows the ROC curves (based on FOI detection) of both detectors. We can see that the matched subspace detector outperforms the reduced-rank Gauss-Gauss detector at virtually all thresholds. The matched subspace detector reduces the false detection rate for FOI by 16.9% (at the 100% detection point) over the Gauss-Gauss detector. Low false detection rate is important in terms of reducing the load on the classifier.

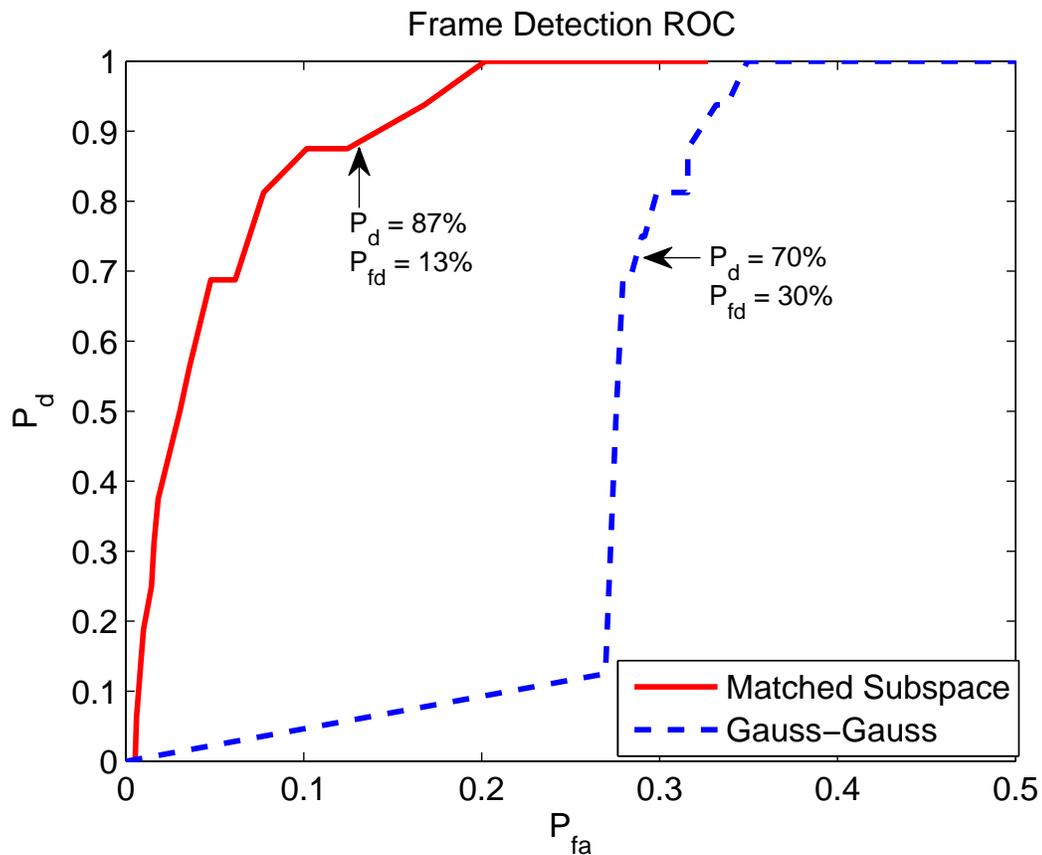


Figure 4.5: ROC for FOI performance comparison of the reduced-rank Gauss-Gauss and matched subspace detectors.

4.4 Conclusion

In this chapter we presented a new detector that casts the detection problem in terms of detecting underlying signal subspace and noise. We showed that using the same block-based framework for FOI, and ROI detection with the matched subspace

detector improved clutter suppression over that of the reduced-rank Gauss-Gauss detector in Chapter 3. This, considerably reduces the load off the classifier. We designed the detector by defining matrix \mathbf{H} and applying SVD on the data matrix \mathbf{T} to give a representative set of basis vectors for mine-like objects. The likelihood map for the matched subspace detector has considerably better contrast when compared to the reduced-rank Gauss-Gauss detector. Additionally the reduced-rank Gauss-Gauss detector achieves a $P_d = 100\%$ detection rate with a $P_{fd} = 36.9\%$ false detection rate, while the matched subspace detector achieves a $P_d = 100\%$ detection rate while maintaining a $P_{fd} = 20.2\%$ false alarm rate. In Chapter 5 we present shape-based classification methods which are applied to both the reduced-rank Gauss-Gauss, and matched subspace outputs after the detection process is completed.

CHAPTER 5

FEATURE EXTRACTION IN ELECTRO-OPTICAL IMAGERY

5.1 Introduction

Once the FOI's are identified a feature extraction process must be used to extract shape-dependent features. Many different shape dependent feature extraction methods exist that can successfully be applied to a wide variety of pattern recognition problems. However, moment-based schemes [27]- [30] are among the most widely used methods as they provide translation, rotation and scaling invariant features ideal for 2-D as well as 3-D pattern recognition applications. In [30], a comparison was made among several types of moments including regular moments, Legendre moments, Zernike moments, and complex moments. These methods were compared in terms of their image representation ability, noise sensitivity, and information redundancy on several character recognition examples. Owing to the fact that the regular moments do not provide an orthogonal representation, the extracted features using this scheme lack optimality in representation. This is in contrast to the orthogonal moments, e.g. Legendre and Zernike moments [27]- [30]. The experiments conducted in [30] indicated that the classification results of the Zernike moments are substantially less sensitive to additive noise effects in the images when compared to the other types. In [28], a similar study was carried out where the regular moments and Zernike moments were used for feature extraction and a back-propagation neural network (BPNN) [19] was employed as a classifier. Due to these useful properties of the Zernike moments we have chosen to use this method for shape-dependent feature

extraction of detected ROI.

In this chapter we show how Zernike moments are used in conjunction with a BPNN classifier in order to classify ROIs from detected FOI as mine-like objects, or background anomalies. Zernike moments are extracted from each ROI in order to characterize the shape of the ROI. Decision between mine-like object, and background clutter is made based on the shape of the ROI in question using the BPNN classifier. Each detected FOI can yield one or more ROIs, while each ROI needs to be classified as either a mine-like object or background clutter. If all detected object in a particular FOI are identified to be background clutter that particular frame will no longer be considered.

This Chapter is organized as follows. In Section 5.2 we review the feature extraction process using the Zernike moments used in this research. Then, a method for reducing the size of the feature sets using Fisher distance measures [37] is described in Section 5.2.2. Next, in Section 5.3 we present the ROI classification process using the reduced feature sets of Zernike moments and a BPNN classifier. The results of this object classification scheme are presented in Section 5.3. Finally concluding remarks are given in Section 5.5.

5.2 Shape-Dependent Feature Extraction

5.2.1 Zernike Moments

Zernike moments are obtained using a complete set of complex polynomials defined in the interior of the unit circle. For an image $f(x, y)$, the Zernike moments of order n with repetition m where $|m| \leq n$ and $n - |m|$ constrained to an even number, are given by

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x, y) V_{nm}^*(\rho, \theta) \quad (5.1)$$

where $x^2 + y^2 \leq 1$ i.e. confined to the interior of the unit circle, ρ is the length of a vector from the origin of the unit circle to (x, y) point, θ is the corresponding phase

angle, and $V_{nm}(\rho, \theta)$ form a complete set of orthogonal complex polynomials over the unit circle. These are defined by

$$V_{nm}(\rho, \theta) = R_{nm}(\rho)e^{jm\theta} \quad (5.2)$$

where the radial function $R_{nm}(\rho)$ is given by

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} \frac{(-1)^s [(n-s)!] \rho^{n-2s}}{s! (\frac{n+|m|}{2} - s)! (\frac{n-|m|}{2} - s)!} \quad (5.3)$$

Note that we have $R_{n,-m}(\rho) = R_{nm}(\rho)$ and also $A_{nm}^* = A_{n,-m}$. The magnitude of the Zernike moments are rotation invariant features. This is due to the fact that upon rotation the complex Zernike moments only acquire a phase factor which does not affect the magnitude of the moments. To make the moments translation invariant, the image is transformed to a new coordinate system by moving the origin to the centroid prior to the moment calculation. Then, the first-order moments of the new image becomes zero. To achieve scale invariance, the image is resized by changing its zeroth-order moment to a predetermined value. A general mapping of type $f(x, y) = g(\bar{x} + \frac{x}{a}, \bar{y} + \frac{y}{a})$ where (\bar{x}, \bar{y}) are coordinates of the centroid of the original image $g(x, y)$ and a is a scaling parameter, is applied prior to computing the Zernike moments. The scale and translation normalization processes affect two Zernike features namely $|A_{00}|$ and $|A_{11}|$. However, $|A_{00}|$ will be the same for all the images and $|A_{11}|$ will be zero. As a result, these moments will not be included in the extracted feature vector and the selected features start from the second-order moments. The great benefit of these moments is their optimality property which is useful when they are used for object classification. In addition, it has been shown that the Zernike moments [28, 27] are more immune to noise and distortion than the regular moments.

For this study Zernike moments up to order $n = 15$ with $m = (0, 2, 4..n)$ for even n , and $m = (1, 3, 5..n)$ for odd n were used to extract features from the detected ROI. This generates a feature vectors of dimension 72 for each ROI. Feature sets of these Zernike moments were built using mine-like object and background clutter

ROI's extracted from training runs outside the runs in Table 2.1. These training runs include: SAM001_004, SAM002_008 (also used to train the detector), and new runs SAM005_003, and SAM005_004. These four runs contain typical target ROI's having regular-shape silhouettes. The feature set built using the above mentioned runs consists of 21 target ROI's, and 21 randomly selected non-target ROI's. The features from these 42 ROI's are extracted and compiled into a feature set, which is then used as the training set for the BPNN classifier. Several mine-like objects and background clutter ROI's from the BPNN training set are shown in Figures 5.1(a) and 5.1(b), respectively.

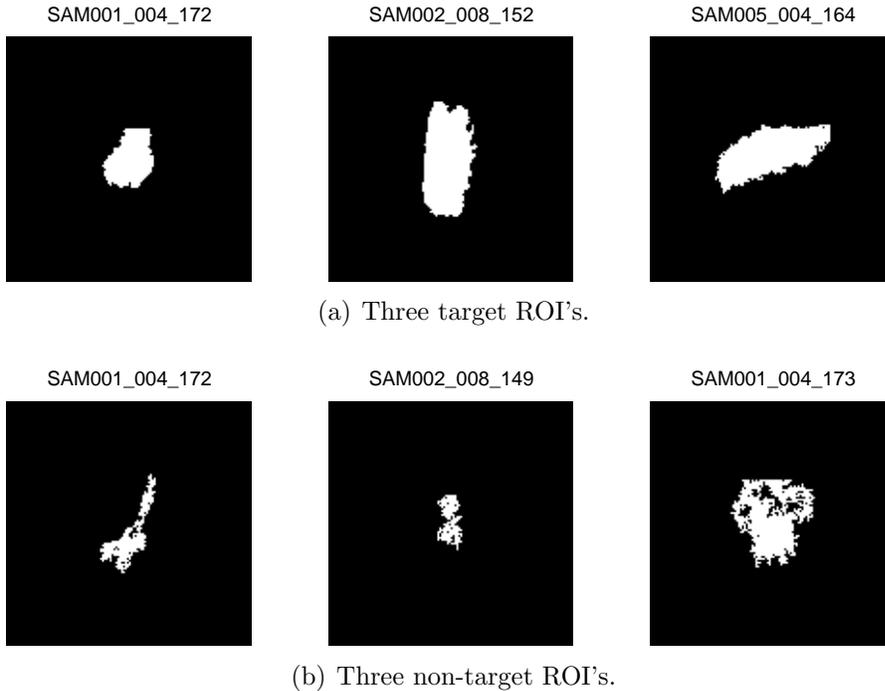


Figure 5.1: Several target and non-target ROI's used for the BPNN training set.

5.2.2 Feature Space Dimensionality Reduction

Since this 72×1 feature vector is of rather large size we employed a distance measure to reduce the size of the feature set by choosing only the top most discriminatory features. To compute the most discriminatory features we compute the following

Fisher distance measure [37] between every pair of features in the target (T) and non-target (NT) training subsets;

$$D_n = \frac{|\mu_{n|NT} - \mu_{n|T}|}{\sigma_{n|NT} + \sigma_{n|T}} \quad (5.4)$$

where D_n is the Fisher distance measure computed for the n th feature. In (5.4) $\mu_{n|T}$, and $\sigma_{n|T}$ are the mean of the n th feature for the target subset, and similarly for $\mu_{n|NT}$, and $\sigma_{n|NT}$. The higher the value of D_n the more discriminatory the n th feature for target vs non-target classification.

To reduce the feature space we first compute D_n for all $n \in [1, 72]$. Next, we order the D_n in descending order, and normalize all D_n to the highest distance found. All features that have normalized distance greater than 0.3 are then kept. This process yields a reduced-dimensional feature space of size 36 which contains the dominant Zernike moments as far as discrimination is concerned. The mean and standard deviation of the dominant features of the BPNN training set are plotted in Figure 5.4(a) for target and background training ROI's.

We can see from Figure 5.4(a) that there is separation between the two reduced dimensional feature sets. We can see that target features exhibit stronger peaks around several features than those of the non-target features. This separation has been found to be adequate for classification of the detected ROIs in this study.

5.3 ROI Classification

In this section, we will study and benchmark the performance of a classifier on the EO-CCD imagery database given in Table 2.1. The classifier considered for this project is a feed forward multi-layer BPNN [19], which uses an error back-propagation method [19], to update the network's connection weights in batch or iterative modes. The whole process to update the connections weights involves applying the data to

the input of the BPNN and propagating it, through the hidden layers and their non-linear activation functions, in the forward direction to generate the outputs. The outputs are then compared to some desired outputs (01 and 10 for target and non-target respectively) that represent the class of the data samples and the errors are back-propagated through the weights of the output and hidden layers, in a consecutive fashion, to generate desired values for updating the hidden layer weights. The objective function for finding the optimum weights is the average mean squared error (MSE) at the output layer. There are several algorithms [19] that can be used to minimize this objective function including the Resilient Back propagation, Conjugate Gradient, Quasi-Newton and Levenberg-Marquard methods.

In BPNN, the choices of learning algorithm, initial values of network parameters, network topology i.e. number of hidden layers and number of neurons in each hidden layer as well as the type of activation function play dominant roles in the overall performance. These choices are typically problem dependent and are made experimentally by trying several structures, topologies and choices of initializations in order to yield the optimally trained network. Additionally, perhaps the most important requirement is the choice of an appropriate and representative set of features for training and testing of the BPNN classifiers. We designate a training set from 21 target samples, and 21 non-target samples (42 total) from runs SAM001_004, SAM002_008, SAM005_003, and SAM005_004 in order to train and validate the BPNN. The network structure was 36 inputs, 36 first hidden layer neurons, 72 second layer hidden layer neurons, and 2 output neurons. The neurons in the hidden layers and output layer had sigmoidal activation functions.

The network is trained using the Levenberg-Marquardt training [19]. Training stops when any of the following conditions occur:

1. The specified maximum number of epochs: EPOCHS is reached.

2. The specified maximum amount of TIME has been exceeded.
3. Performance GOAL has been reached.
4. The performance gradient falls below MINGRAD.
5. The learning rate MU exceeds MU_MAX.
6. Performance on the validation set has increased more than MAX_FAIL times since the last time it decreased (when using validation).

These training parameters are outlined in Table 5.1.

Table 5.1: MATLAB Neural Network Training Parameters

<code>net.trainparam.param</code>	Value
<code>epochs</code>	25
<code>time</code>	Inf
<code>goal</code>	0
<code>max_fail</code>	6
<code>min_grad</code>	1.0000e-010
<code>mu_max</code>	1.0000e+010

The initialization function used to initialize the layer’s weights and biases is the Nguyen-Widrow initialization algorithm [19]. This algorithm chooses values in order to distribute the active region of each neuron in the layer evenly across the layer’s input space. The MSE is used as the network’s performance measurement function, and the 42 training samples are divided randomly 50/50 between training, and validation sets for training the network, and identifying the best trained network. The purpose of validation is to check for the presence of over-training and optimally select parameters in order to minimize the error. The validation set is formed by randomly selecting 50% (half target, half non-target) of the 21 target samples, and 21 non-target samples (42 total) from runs SAM001.004, SAM002.008, SAM005.003, and SAM005.004. Throughout this study we have held network parameters at their default values within MATLAB.

5.4 Classification Results and Observations

5.4.1 ROI Classification Results

In this section we present the performance of the classifier system on the testing set in Table 2.1. The ROI detection results have already been presented in Chapters 3 and 4. Here we have a two-class ROI prediction problem, in which the outcomes are labeled either as positive (P) or negative (N) classes. There are four possible outcomes from a two-class classifier. If the classifier outcome is P and the actual value is also P, then it is called a true positive (TP); while if the actual value is N then it is said to be a false positive (FP). Conversely, a true negative (TN) is occurred when both the predicted outcome and the actual are both N, and false negative (FN) is when the predicted outcome is N while the actual is P. Based upon these, we summarize the following performance measures associated with the ROI classification.

1. True Positive Rate (TPR) = $TP/(TP+FN)$
2. False Positive Rate (FPR) = $FP/(FP+TN)$

The confusion matrices for the classifier using the matched subspace and reduced-rank Gauss-Gauss detectors ROI outputs are shown in Tables 5.2 and 5.3. We can see that the classifier in the matched subspace case classifies far less objects than that of the reduced-rank Gauss-Gauss. Since the network is trained with only 21 hand picked target and non-target ROI's the network is not able to cover the wide range of non-target ROI scenarios encountered in the testing set. This can be solved by using a classifier equipped to handle data imbalance between target and non-target samples. An example of such a classifier is support vector machines (SVM) [19].

Figure 5.2 shows the ROC curve for ROI classification. We can see that the classifier performs worse on the ROI's generated using matched subspace detector when comparing to those of the reduced-rank Gauss-Gauss detector's output. This result is somewhat misleading since the classifier in the matched subspace case operates on

Table 5.2: Confusion Matrix for Matched Subspace Output

	truth	P	N	Total
predicted				
P		16	0	16
N		302	123	425
Total		318	123	

Table 5.3: Confusion Matrix for Reduced-Rank Gauss-Gauss Output

	truth	P	N	Total
predicted				
P		16	0	16
N		363	419	782
Total		379	419	

far fewer ROIs than in the reduced rank Gauss-Gauss case. The classifier operating on the matched subspace detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm $P_{fa} = 71\%$. The classifier operating on the reduced-rank Gauss-Gauss detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm $P_{fa} = 46\%$. The classifier handles less ROI's in the case of the matched subspace detector, and since the BPNN is trained with a balanced training set the classifier does not properly handle the large number of non-target ROI's encountered in the testing set.

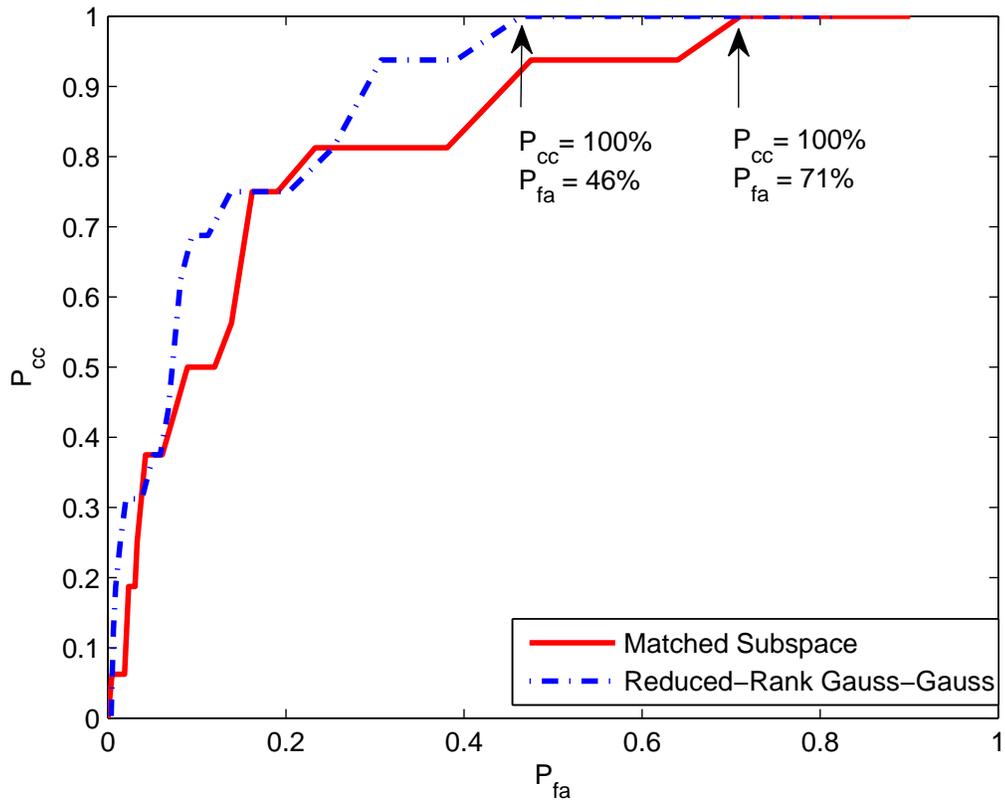


Figure 5.2: ROI classification performance receiver operating characteristics (ROC) for matched subspace detector, and reduced-rank Gauss-Gauss detector.

5.4.2 Overall FOI Detection/Classification Results

We also look at the combined detector/classifier in terms of FOI performance. We look to see how many FOI have been eliminated after the classifier has been invoked. We look at the FOI performance due to the fact that the operator will evaluate the output of the system by looking at FOI, and ROI within the FOI. The frame reduction performance of the combined detector and classifier system for matched subspace and reduced-rank Gauss-Gauss are presented in Tables 5.4 and 5.5, respectively. The columns of Tables 5.4 and 5.5 are summarized by following:

1. FA FOI: The number of FOI false alarms per run using the stand alone detectors (e.g. matched subspace and reduced-rank Gauss-Gauss detector alone).
2. FA FOI (w/ BPNN): The number of false FOI per run after extracting features and applying runs to the BPNN classifier.
3. FOI Reduction %: The percentage false FOI reduced per run by using the BPNN classifier

From Tables 5.4 and 5.5 we see that the BPNN reduced detected false FOI by 3.3% when using the matched subspace detector, while the BPNN significantly reduced detected false FOI by 15.1% when using the reduced-rank Gauss-Gauss Detector. Along with Tables 5.4 and 5.5.

Figure 5.3 presents the overall frame detection ROC (FOI-ROC) for both matched subspace, and reduced-rank Gauss-Gauss detectors. This figure shows the FOI ROC for the stand alone detectors, as well as false alarm suppression achieved using the BPNN. Upon evaluating the performance of the overall system on data of Table 2.1, it was found that the matched subspace detector has a detection rate of $P_d = 100\%$ at the cost of probability of false detection $P_{fd} = 20.2\%$. Features extracted from the matched subspace detector's output passed to a BPNN classifier yields a true FOI

Table 5.4: Detection Results for Matched Subspace Detection and Classification

Run	FA FOI	FA FOI(w/ BPNN)	FOI Reduction %
SAM001_003	3/42	3/42	0.0%
SAM004_001	0/35	0/35	0.0%
SAM22_011	9/35	7/35	5.7%
SAM23_003	60/293	52/293	2.7%
SAM23_004	62/287	52/287	3.4%
SAM23_005	63/293	57/293	2.0%
TargetY8_001	31/136	27/136	2.9%
TargetY8_003	9/29	6/29	10.3%
TargetY8_004	1/32	0/32	3.1%
TargetY8_006	28/135	19/135	6.6%
Totals	266/1317	223/1317	43/1317
Percentage	20.2%	16.9%	3.3%

Table 5.5: Detection Results for Reduced-Rank Gauss-Gauss Detection and Classification

Run	FA FOI	FA FOI(w/ BPNN)	FOI Reduction %
SAM001_003	2/42	1/42	2.3%
SAM004_001	7/35	4/35	8.5%
SAM22_011	12/35	10/35	5.7%
SAM23_003	142/293	73/293	23.5%
SAM23_004	92/287	75/287	5.9%
SAM23_005	129/293	66/293	21.5%
TargetY8_001	47/136	34/136	9.5%
TargetY8_003	11/29	2/29	31.3%
TargetY8_004	16/32	0/32	50.0%
TargetY8_006	28/135	22/135	4.4%
Totals	486/1317	287/1317	199/1317
Percentage	36.9%	21.7%	15.1%

detection rate of $P_d = 100\%$ at the cost of $P_{fd} = 16.8\%$ hence reducing the detected false FOI by 3.3%. The reduced-rank Gauss-Gauss detector has a detection rate of $P_d = 100\%$ at the cost of probability of false detection $P_{fd} = 36.9\%$, while features extracted from the reduced-rank Gauss-Gauss detector's output passed to a BPNN classifier yields a true FOI detection rate of $P_d = 100\%$ at the cost of $P_{fd} = 21.7\%$ hence significantly reducing the detected false FOI by 15.1%.

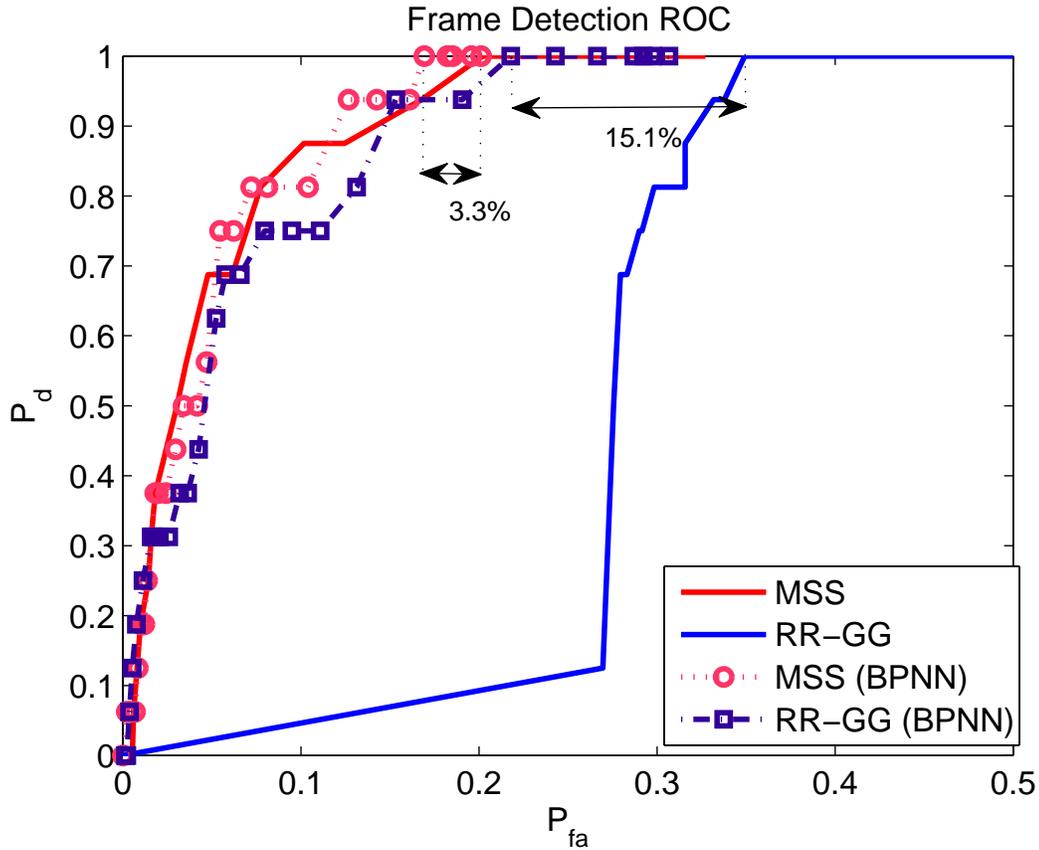


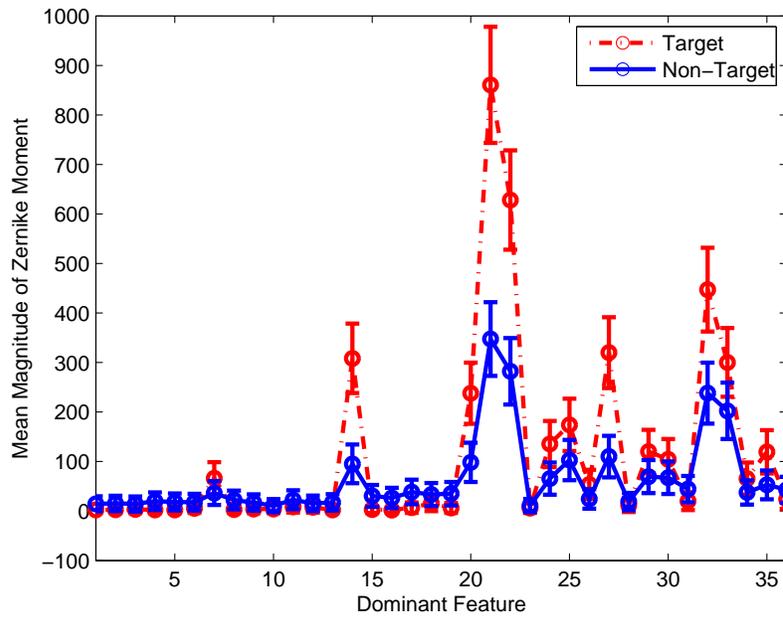
Figure 5.3: FOI performance receiver operating characteristics (ROC), with and without BPNN.

It is notable that the number of frames has been reduced, and the matched subspace detector-classifier finds 302 false ROIs, while the reduced-rank Gauss-Gauss detector-classifier finds 366 false ROI's. To look into the number of false ROIs detected we generate the separation between the two reduced dimensional feature sets

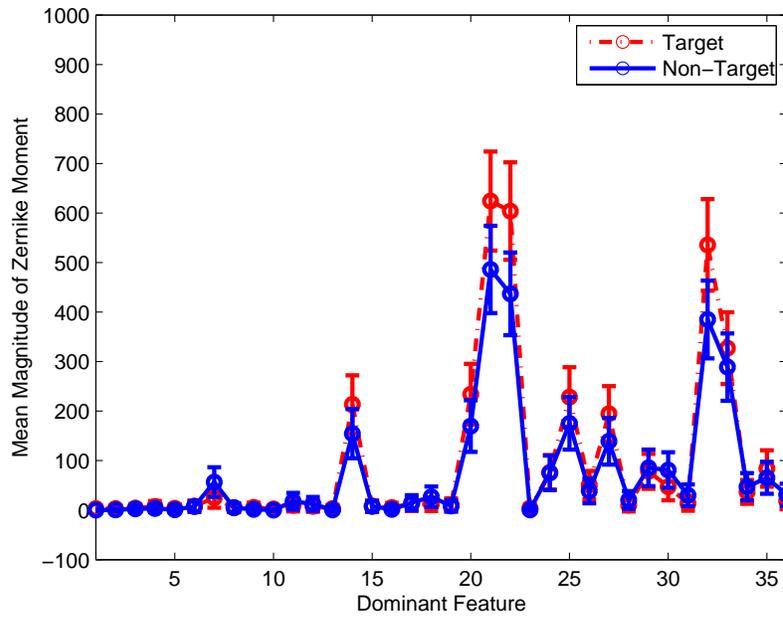
which have been detected, and classified using the matched subspace detector. Figure 5.4(b) shows a similar plot as in Figure 5.4(a) for the testing set as can be seen the shape based features in the testing set overlap with those of the targets much more than in the training set (Figure 5.4(a)), this also contributes to the large number of false ROI's detected in this testing set.

Figures 5.5 and 5.6 show examples of misclassified target and non-target ROI's (using the threshold value at the ROC knee points). Figure 5.5 shows examples of false negative objects (Likelihood map and centered detected ROI). The detected ROI's in Figure 5.5 are centered for computation of Zernike moments. One main reason for false negatives is fact that partial targets can exhibit features which resemble small background anomalies. Notice all examples of false negative objects in Figure 5.5 are ROI's of partial targets. Figure 5.6 shows example of false positive objects (Likelihood map and centered detected ROI). False positives can occur when detected background anomalies ROI's are 'filled in' (solid) and resemble an ROI of a mine-like object. As mentioned before the main problem, is that impedes the classification process the overwhelming number of non-target anomalies with similar features as those of partial targets.

While these results are promising they highly depend on the primary detector's (e.g. matched subspace or reduced-rank Guass-Gauss) choice of an appropriate threshold for robust silhouette definition. If the primary detector fails to output complete silhouettes of detected mine-like or non-mine-like objects subsequent feature extraction, and classification become difficult, and may eventually fail.



(a) Plot of mean and standard deviation for all classified object Zernike features in the reduced ROI training feature set detected using the matched subspace detector.



(b) Plot of mean and standard deviation for all classified object Zernike features in the reduced ROI testing feature set detected using the matched subspace detector.

Figure 5.4: Plot of mean and standard deviation for Zernike features in the reduced feature sets

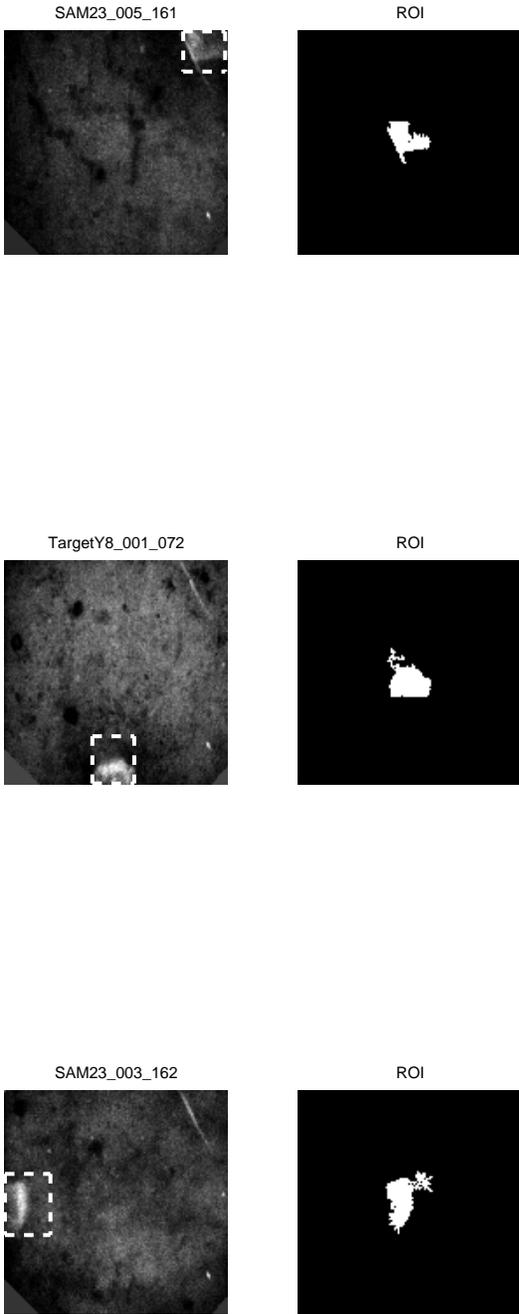


Figure 5.5: Example of false negative (ROI's). The likelihood map output from matched subspace detector (left), and false negative ROIs (right).

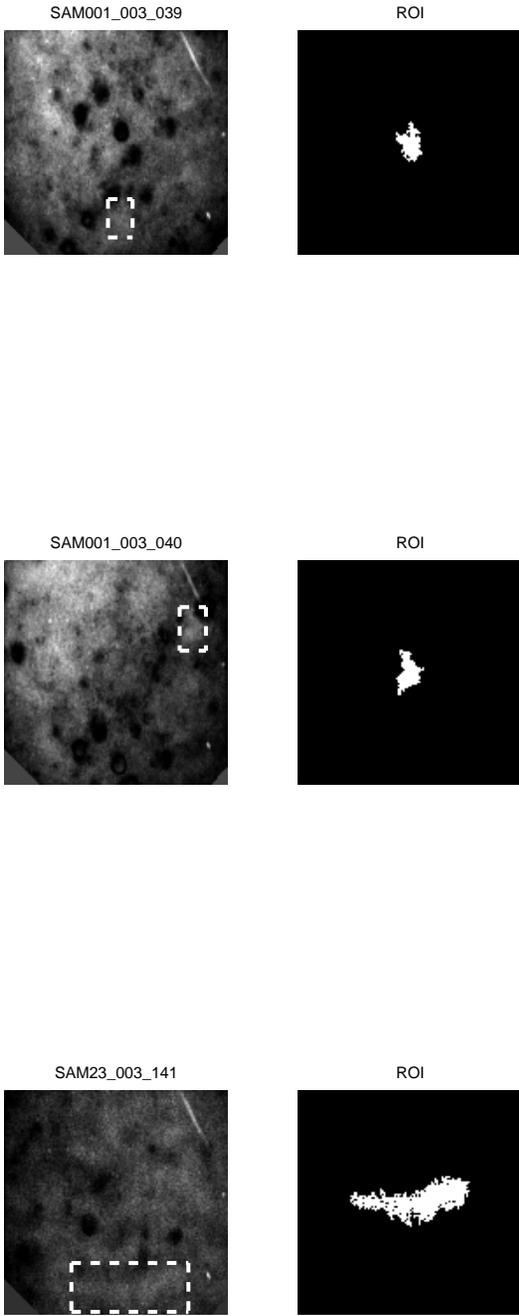


Figure 5.6: Example of false positive (ROI's). The likelihood map output from matched subspace detector (left), and false negative ROIs (right).

5.5 Conclusion

In this chapter we have shown how Zernike moments can be used in conjunction with a BPNN classifier to classify ROIs from detected FOI as mine-like objects or background anomalies. Zernike moments are extracted from each ROI in order to characterize the shape of the detected ROI. Classification of the mine-like object, and background clutter is decided based upon shape-based features extracted from ROIs in question. A method for reducing the dimension of the feature space using Fisher distance measures was also discussed that allows selection of the most discriminatory features for this two-class classification problem. The process of ROI classification using this reduced dimensional feature set of Zernike moments and a BPNN classifier was then discussed.

The classifier operating on the matched subspace detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm $P_{fa} = 71\%$. The classifier operating on the reduced-rank Gauss-Gauss detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm $P_{fa} = 46\%$. However, the former case classifier handles less ROIs than in the latter. On the other hand since the BPNN is trained with a balanced training set the classifier cannot successfully classify the large number of non-target ROIs encountered in the testing set. This can be overcome by using a classifier which can handle data imbalance issues, e.g. SVN [19].

The frame reduction capability of the combined detector-classifier system was also demonstrated. The stand alone matched subspace detector achieves $P_{fd} = 20.2\%$, while the combine system achieves $P_{fd} = 16.8\%$, the combine system reduces the false detected FOI by 3.3% over that of the stand alone detector. The stand alone reduced-rank Gauss-Gauss detector achieves $P_{fd} = 36.9\%$, while the combine system achieves $P_{fd} = 21.7\%$, the combine system reduces the false detected FOI by 15.1% over that of the stand alone detector. All FOIs can be successfully identified, and

target ROIs subsequently classified correctly using the methods implemented in this research. It has been found that false positives and false negative misclassifications occur when partial mine-like object ROIs are detected, or when small background anomalies appear solid. The main problem is the overwhelming number of non-target anomalies with similar features as those of partial targets. Very promising results have been achieved, but room for improvement still exists in terms of optimizing classifier paradigm and training.

CHAPTER 6

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

This research project involved the development of an underwater target detection and classification system for electro-optical EO-CCD imagery data. Automatic recognition of underwater objects, such as mine-like objects, from CCD imagery poses many technical problems. The targets in the EO-CCD imagery could be proud or partially buried making the contrast and texture of mine-like objects and background clutter very similar. The CCD sensor generates a sequence of snapshots taken from ocean bottom as opposed to the STIL sensor, which used a LIDAR to scan an area to generate contrast and range maps after the data rendering process [1] - [12]. After analyzing the data it was found that different target detection and segmentation methods are needed than those developed for the STIL imagery in [1], [4].

This work has led to successful development and implementation of an automatic target recognition system comprising a block-based detector and ROI segmentation, Zernike-based feature extractor and a BPNN based classifier for the new electro-optical CCD imagery data. The detector takes a sequence of frames as an input and produces detected frames (FOIs) numbers, object locations, and confidence scores associated with each detected ROI as outputs. The detector also generates likelihood maps indicating the location of each detected ROI as well as the binary silhouette of each segmented object.

An application GUI has been developed (See Appendix A) which allows one to study the performance of different detection systems. The GUI allows a user to build

training sets for the detectors, analyze likelihood maps, segment ROI's, and build Zernike-based feature sets for a classifier all interactively.

The block-based detectors developed here have been shown to generate likelihood maps with relatively high separation between mine-like object and background blocks. The two-step process which primarily identifies FOI in a data run and then extracts ROI within the FOI has been proven to be successful on the EO imagery data considered in this work. The Gauss-Gauss detector's rank-reduction method significantly improved the definition of the likelihood maps generated by the detector. The matched subspace detector is formed using an SVD on a preselected training data matrix. The SVD generates optimal basis vectors used to generate a signal subspace which is used to project given samples and measure how those samples align with the subspace. The matched subspace detector improves clutter suppression and generation of complete silhouettes in segmented ROI's. Both detectors remove, on average, more than half the frames on the tested data with matched subspace having $P_d = 100\%$ and $P_{fd} = 20.2\%$, and reduced-rank Gauss-Gauss having $P_d = 100\%$ and $P_{fd} = 36.9\%$.

We have shown that ROI can be successfully segmented from the input frames, and shape-based Zernike moments can be successfully used to identify a detected ROI as a mine-like objects or background clutter using a neural network-based classifier. In this work we have employed robust feature extraction methods as well as implemented a distance measure to reduce the dimensionality of the feature space of Zernike moments used for ROI classification. The reduced feature sets extracted from the detected ROI's have been found to be separable for mine-like object discrimination from background clutter.

The classifier operating on the match subspace detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm

$P_{fa} = 71\%$. The classifier operating on the reduced-rank Gauss-Gauss detector's output achieves probability of correct classification $P_{cc} = 100\%$ at the cost of probability of false alarm $P_{fa} = 46\%$. The classifier handles less ROIs in the case of the matched subspace detector, and since the BPNN is trained with a balanced training set the classifier does not properly handle the large number of non-target ROIs encountered in the testing set. This can be handled by using a implementing a classifier which is trained on a much larger set of non-target ROI scenarios.

It has been found that the use of BPNN classifier significantly reduces false FOI detections by 15.1% for the reduced-rank Gauss-Gauss detector, and 3.3% for the matched subspace detector. Using the detector and classifier together yeilds $P_d = 100\%$ with $P_{fd} = 16.9\%$ for the matched subspace detector and BPNN, and $P_d = 100\%$ with $P_{fd} = 21.7\%$ for the reduced-rank Gauss-Gauss detector and BPNN.

Overall, the algorithms developed in this research provided promising results for underwater target detection and classification from CCD imagery. Among the desirable characteristics of the proposed methods is the simplicity of the detection algorithms for detecting FOI in a run, while automatically segmenting the ROI of mine-like objects. Unlike the methods in [1], no preprocessing is needed here and one algorithm is used for both FOI identification and ROI extraction. While the system developed here focuses on the detection and classification of mine-like objects, it is certainly applicable to other areas such as image-based fault detection and quality control.

6.1 Future Work

Although, the block-based detectors developed in this thesis offer powerful tools for detection of underwater mine-like objects, there are several important areas and extensions that can be pursued in the future. These include, but are not limited to:

- 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 block-based detector would be to test the performance on more data to prove the usefulness of the detection systems developed in this thesis in different environmental and operating conditions. The testing on more difficult data sets provided by AST or NSWCC as well those including more man-made non-targets should also be done in the future. More specifically, a study on the effect of different bottom types, target orientations, CCD resolution, and SNR on the probability of detection and false alarm rates would be insightful and help to illustrate the real effectiveness of the detector for realistic underwater target detection problems.
- The main development in this thesis was devoted to the detection system design for underwater target detection from EO imagery, while a small portion of the work was carried out on the classification. Potential extension of this research would be to study the use of other features for classification of targets and non-targets, and background clutter. Additionally, different classification paradigms can be tried to optimize the overall rates.
- The block-based detector developed in this research is applicable not only to EO image detection, but also to other sensory systems which generate images, i.e. magnetic, infrared, and sonar imagery. A study of its usefulness on these types of sensing modalities would highly be valuable.
- Another possible extension to the target detection from EO imagery is multiple hypothesis testing. There has been a large amount of research devoted to multiple hypothesis testing, especially in the areas of bioinformatics, genomics, and brain imaging [38], [39]. The idea behind a multiple hypothesis testing approach is that by performing multiple statistical tests where a number of hypotheses

is rejected and another number of hypotheses is accepted allowing us to identify the detected ROI to belong to a particular type of objects (e.g. mine-like objects versus natural/man-made clutter and background alone). 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 hence reduce the overall false alarm rate.

- This detector can potentially be improved by using the signal plus noise plus interference model as presented in [15] and [40]. Further research using these generalizations may improve detector performance, by designing detectors which detect targets and remove structured interference from the frames.
- Another possibility for research lies in using different methods than SVD to choose an optimal basis for signal and noise basis vectors used in the matched subspace detector. Among possible methods for choosing the best basis [41,42].
- Room for improvement also exists in terms of optimizing the classifier paradigm and training in order to account for the large number of detected ROIs, and data imbalance issues. This could mean using classifiers such as SVM [19].

REFERENCES

- [1] M. R. Azimi-Sadjadi, “Underwater target identification using electro-optical imagery, Colorado State University,” *Final Report to Coastal System Station*, 2003.
- [2] J. Salazar and M. Azimi-Sadjadi, “Identification of underwater mines from electro-optical imagery using an operated-assisted reinforcement on-line learning,” *Proc. OCEANS 2003*, vol. 1, pp. 124–131, April 2003.
- [3] M. P. Strand, “Underwater electro-optic system for mine identification,” *Proceedings for Autonomous Vehicles in Mine Countermeasures Symposium, Monterey, CA*, April 1995.
- [4] G. Tao, M. R. Azimi-Sadjadi, and A. Nevis, “Underwater target identification using GVF snake and Zernike moments,” *Proc. of the MTS/IEEE Oceans*, vol. 3, pp. 1535–1541, Oct 2002.
- [5] A. Gleckler *et al.*, “Streak tube imaging lidar for electro-optic identification,” *Fourth International Symposium on Technology and the Mine Problem, Monterey, CA*, March 2000.
- [6] E. Watson, “New imaging modalities for laser-based systems,” *IEEE Proceedings: Aerospace Conference*, vol. 3, pp. 1593–1599, March 2001.
- [7] M. D. Iwanowski, “Surveillance unmanned underwater vehicle,” *Proc. of OCEANS '94*, vol. 1, pp. 116–119, September 2004.
- [8] J. J. Shirron and T. E. Giddings, “A comprehensive model for performance prediction of electro-optical systems,” *Metron, Inc.*, vol. 1, 2007.

- [9] —, “A model for the simulation of a pulsed laser line scan system,” *Proc. of OCEANS '06*, vol. 1, pp. 1–6, September 2006.
- [10] A. J. Nevis, “Underwater electro-optic system for mine identification,” *Third International Symposium on Technology and the Mine Problem, Monterey, CA*, April 1998.
- [11] A. Nevis, J. S. Taylor, and B. Cordes, “A baseline object detection algorithm using background anomalies for electro-optic identification sensors,” *Proc. of 2002 MTS/IEEE Oceans Conference, Biloxi*, vol. 3, pp. 1546–1554, October 2002.
- [12] J. S. Taylor and M. C. Hulgan, “Electo-optic identification research program,” *Proc. of 2002 MTS/IEEE Oceans Conference, Biloxi*, vol. 2, pp. 994–1002, October 2002.
- [13] R. Manley, “Underwater EO imager for partially buried mines,” *Naval Surface Warfare Center Panama City*, Jan 2008.
- [14] L. L. Scharf and B. D. Van Veen, “Low rank detectors for Gaussian random vectors,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 35, pp. 1579–1582, Nov 1987.
- [15] L. L. Scharf and B. Friedlander, “Matched Subspace Detectors,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 2146–2157, Aug 1994.
- [16] H. L. Van Trees, *Detection, Estimation, and Modulation Theory Part I*. John Wiley and Sons, 1968.
- [17] L. L. Scharf, *Statistical Signal Processing*. Addison-Wesley, 1991.

- [18] R. M. Haralick, K. Shanmugan, and I. H. Dinstein, “Textural features for image classification,” *IEEE Trans. on Syst., Man, Cybern.*, vol. 3, pp. 610–621, March 1973.
- [19] S. Haykin, *Neural Networks: A comprehensive foundation- second ed.* Prentice Hall, 1999.
- [20] C. Cortes and V. Vapnik, *Support Vector Networks Machine Learning.* John Wiley and Sons, 1995.
- [21] (2008, September) Wikipedia, The Free Encyclopedia. [Online]. Available: http://en.wikipedia.org/wiki/Naval_mine
- [22] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions.* Wiley-Interscience, 1996.
- [23] C. Xu and J. L. Prince, “Snakes, Shapes, and Gradient Vector Flow,” *IEEE Trans. on Image Processing*, vol. 7, Mar 1998.
- [24] R. C. Gonzalez and R. E. Woods, *Digital Image Processing.* Prentice Hall, 2001.
- [25] C. Teh and R. Chin, “On image analysis by the methods of moments,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 496–513, July 1988.
- [26] A. Khotanzad and J. Lu, “Classification of invariant image representation using a neural network,” *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 10, pp. 1028–1038, June 1990.
- [27] A. Khotanzad and J. Liou, “Recognition and pose estimation of unoccluded three-dimensional objects from a two-dimensional perspective view by banks of neural networks,” *IEEE Trans. on Neural Networks*, vol. 7, pp. 897–906, July 1996.

- [28] A. Khotanzad and J. Lu, “Classification of invariant image representation using a neural network,” *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 38, pp. 1028–1038, June 1990.
- [29] M. R. Teague, “Image analysis via the general theory of moments,” *Journal of Optical Soc. of Amer.*, vol. 70, pp. 920–930, August 1980.
- [30] C. Teh and R. Chin, “On image analysis by the methods of moments,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 10, pp. 496–513, July 1988.
- [31] M. Azimi-Sadjadi and S. Stricker, “Detection and classification of buried dielectric anomalies using neural networks—further results,” *IEEE Trans. on Instrumentation and Measurement*, vol. 43, pp. 34–39, February 1994.
- [32] (2007, June) Dvc-1500m datasheet. [Online]. Available: <http://www.dvcco.com/PDFs/datasheets/cameras/DVC-1500M.pdf>
- [33] (2005, October) Philips lumiled luxeon flood 18 led illuminator. [Online]. Available: <http://www.philipslumileds.com/pdfs/DS24.PDF>
- [34] (2010, April) Wikipedia, The Free Encyclopedia. [Online]. Available: <http://www.bluefinrobotics.com/bluefin12.htm>
- [35] E. L. Lehman, *Testing Statistical Hypotheses*. New York: Wiley, 1986.
- [36] G. H. Golub and C. F. Van Loan, *Matrix Computations*. The Johns Hopkins University Press, 1996.
- [37] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley and Sons, 2001.

- [38] 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.
- [39] 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.
- [40] L. L. Scharf and R. T. Behrens, “Signal Processing Applications Oblique Projection Operators,” *IEEE Transactions on Signal Processing*, vol. 42, pp. 1413–1423, Aug 1994.
- [41] M. Gharavi-Alkhansari and T. Huang, “A Fast Orthogonal Matching Pursuit Algorithm,” *Proc. of IEEE Acoustics Speech and Signal Processing*, vol. 3, pp. 1389 – 1392, May 1998.
- [42] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, Dec 1993.

APPENDIX A

GRAPHICAL USER INTERFACE (GUI)

A.1 Introduction

In this appendix, we provide a detailed description of the designed Graphical User Interface (GUI) software to study the performance of block-based detection methods on the EO-CCD data set, or any other digital image database. The GUI plays an integral role in evaluating, and building training data sets for different detectors tried in this project, as well as building feature sets to train a neural network-based classifier. This Appendix outlines the features of the GUI, and serves as a tutorial for its use.

The GUI allows a user to interactively select target and background blocks by drawing boxes on a frame to be included in the detector's training set. The GUI also allows the user to measure likelihood values in the generated likelihood map and displays them in a histogram plot below the likelihood map image (this is useful for selecting detector thresholds). Additionally, the GUI allows the user to choose rank of the detector, the type of detector, and compute the likelihood map of a test image based on the selected training set. Once a training set is selected for the detector, the user can measure the likelihood values for target and non-target blocks. The GUI can then segment the likelihood map based on the selected threshold and extract object ROI, and compute the Zernike-based features of those ROI. Finally, feature vectors can be used to form a feature set to train and test a neural network-based classifier.

Another feature of the GUI is that it allows the user to import previously saved training and feature sets, as well as export detector training and feature sets for later use, if needed. This GUI uses the MATLAB programming language and toolboxes,

and provides an easy to use software package allowing users who are not experts in detection or neural networks to explore the methods discussed in this thesis.

A.2 GUI Usage Reference

The GUI implements block-based detection methods for EO-CCD data sets. The overall capabilities of the GUI are: (1) easy and quick comparison of block-based detectors in terms of evaluating and assessing separation between mine-like objects and background clutter; (2) easy design of training sets for use with block-based detectors; (3) add and save multiple training blocks for easy comparison of different training sets; (4) visualize the output of the detector and measure likelihood separation between mine-like objects, non-targets, and background clutter. Figure A-1 shows the main MATLAB figure of the GUI.

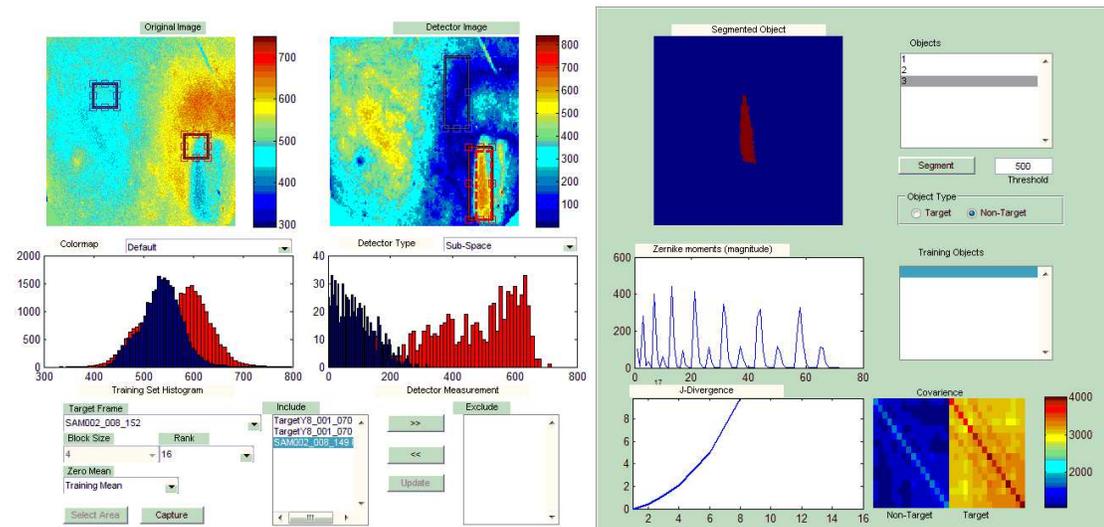


Figure A-1: GUI tool used for building training sets for the detector, and feature sets for the classifier.

The following walks through typical usage of the GUI. Figure A-1 shows the main windows of the GUI which is mainly used to build training sets for a detector, and apply the trained detector on test images in order to rapidly evaluate the detector. Tables A.1 A.2 A.3 serve as a function reference for using the GUI. In this appendix

GUI commands, controls, displays, and filenames are typed in Typewriter text. The first task when using the GUI is to load images (a run) into the GUI from either `ALL_TARGETS_STRUC.mat` or `<run_name>_<frame#>.mat`. This is accomplished using the File menu.

The following steps outline how to use the GUI to build a new training set for the detector, and evaluate the trained detector on a test image. In order to use the block-based likelihood detector a ‘training set’ must be selected. The selection of a set of blocks from targets and background is required in order to compute the covariance matrices associated with H_0 and H_1 for Gauss-Gauss and to build and form \mathbf{H} , and $\mathbf{P}_{\mathbf{H}}$ for matched subspace. This process is subjective in that the blocks used for the training must be hand picked from frames which are believed to represent a wide range of target scenarios.

1. Load Targets using the the File menu `File -> Load Targets` or `File -> Open`
2. Choose desired frame from the `Target Frame` Drop down menu.
3. Adjust detector parameters such as `Block size`, `Rank`, `Detector Type`, `Zero Mean`, etc.
4. Use the `Select Area` button located on the southwest corner of the GUI to bring up the H_0 (blue=non-target) and H_1 (red=target) draggable-resizable selection boxes within the `Original Image` panel.
5. Once the selection boxes are positioned use the `Capture` button to include the data snippet in the current training set.
6. Use the `Update` button to evaluate the detector on the current (or desired) `Target Frame`.

7. Capture more data snippets by repeating steps 4, 5, and 6, while using the >> button to exclude undesired data snippets.
8. Use the Red and Blue draggable-resizable selection boxes located in the `Detector Image` panel to measure the separation between target and non-target, using `Detector Measurement` histogram.
9. Once satisfied with the training set export the training set using `File -> Export Training Set`.

Once a training set is built and the user is satisfied with the detector separation between H_0 and H_1 , the user is ready to move on to building a feature set for the classifier. The following step walk through how to build a shape based feature set for a classifier based on Zernike moments.

1. Import (using `File -> Import Training Set`) or build a new detector training set.
2. Load Targets using the the File menu `File -> Load Targets` or `File -> Open`.
3. Choose desired frame from the `Target Frame` Drop down menu.
4. Use the `Update` button to evaluate the detector on the current (or desired) `Target Frame`.
5. Use the Red and Blue draggable-resizable selection boxes located in the `Detector Image` panel to decide on a value for the `Threshold` numeric box.
6. Use the `Segment` button to segment all objects above `Threshold` in the `Detector Image`, and populate the `Objects` listbox.

7. Designate objects to be in the classifier training set as either target, or non-target using the `Object-Type` radio buttons, and then double clicking the desired object in the `Objects` listbox.
8. Repeat steps 3-7 until the desired number of classifier training objects have been selected.
9. Once satisfied with the classifier training set export the training set using `File -> Export Feature Set`.

A.3 GUI Function Reference

Table A.1: File Menu

Item	Description
Load Targets	Loads targets from <code>ALL_TARGETS_STRUC.mat</code> and displays them in the <code>Original Image</code> panel
Open...	File Open dialog for choosing <code><run_name>.<frame#>.mat</code> files
Print...	Prints the entire MATLAB GUI figure
Import Training Set	Imports previously exported detector training set from <code>*.txr</code> files
Export Training Set	Exports detector training set to <code>*.txr</code> file
Export Feature Set	Exports zernike feature set to <code>*.fxr</code> file
Close	Closes the GUI window

Table A.2: GUI Plots and Graphical Data Display

Item	Type	Description
Original Image	Image	displays a loaded original image selected from Target Frame drop down
Detector Image	Image	displays detector output
Training Set Histogram	Histogram	shows pixel intensity value distribution for target and non-target training set
Detector Measurement	Histogram	shows likelihood intensity value distribution for Red(target) and Blue(non-target) boxes
J-Divergence	Graph	Real time plot of J-Divergence as a function of rank
Zernike Moments	Graph	Plot of The magnitude of Zernike moments up to 15th order
Segmented Object	Image	silhouette of currently selected detected object
Covariance	Image	Plots detector multivariate covariance for target and non-target training sets

Table A.3: GUI Controls

Item	Type	Description
Target Frame	Drop Down Menu	displays a loaded original image selected from Target Frame drop down
Block Size	Drop Down Menu	Used to select the training set and detector block size
Rank	Drop Down Menu	Used for selecting the detector rank (Log-Likelihood/Gauss-gauss only)
Zero Mean	Drop Down Menu	how to zero mean each frame ;Training set mean— Run mean;
Colormap	Drop Down Menu	Choose the color map of all images
Detector Type	Drop down Menu	;Log-Likelihood—Matched-Subspace;
<< (Include)	Button	Include selected (excluded) data snippet in training set
>> (Exclude)	Button	Exclude selected (included) data snippet in training set
Select Area	Button	Generate area selection box in the original image panel for training set block selection (Red=Target, Blue=Non-Target)
Capture	Button	Capture current Red(target) and Blue(non-target) selection boxes and includes them in the current training set
Update	Button	Run the detector based on the current image (based on current (include) training set), and generate new detector image
Objects	Listbox	List of detected objects at the given threshold. Each object is double clickable to designate it as a target or non-target in the feature set using the Object-Type radio button.
Segment	Button	Segment objects from current detector image using current threshold numeric box, and populate Objects listbox
Threshold	Numeric textbox	Numeric value to threshold detector image.
Object-Type	Radio Button	Used to designate selected object in Objects list box as target or non-target for object feature set
Training Objects	Listbox	Objects currently in the neural network feature set