NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

ACOUSTIC IMAGE MODELS FOR NAVIGATION WITH FORWARD-LOOKING SONARS

by

Theodore D. Masek Jr.

December 2008

Thesis Advisor:	Mathias Kölsch
Second Reader:	Kevin Squire

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 2008	Master's Thesis

4. TITLE AND SUBTITLE: Acoustic Image Models for Navigation with Forward-Looking Sonars

5. FUNDING NUMBERS

6. AUTHOR(S): Masek, Theodore

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
N/A

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES: The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)
Cost and miniaturization of autonomous unmanned vehicles (AUV) drive component reuse and better sensor data analysis. One such component is the forward looking sonar (FLS) which can be used for obstacle avoidance and to extract vehicle state information. However, autonomous feature extraction of images from the FLS is difficult due to the noise inherent in the sensor and the sensor's susceptibility to interference from other acoustic devices. This thesis investigated techniques to detect and classify common acoustic noise artifacts and common objects in a single frame. Other techniques require three or more frames to filter objects from other noise sources. A combination of probabilistic and template-based models were used to successfully detect and classify acoustic noise and objects. One common noise source is the micro modem which was detected 100% of the time with 1% false positives. Objects such as the ocean floor were correctly classified more than 93% of the time in most sites. Due to the short development time frame, the software was developed with a two-stage approach. First, a high level scripting language was used for rapid prototyping of different classification techniques. In order to meet the time-constrained requirements of the target software, the classification algorithms were encapsulated as C++ classes in an object oriented design once the desired techniques were identified.

14. SUBJECT TERMS
Computer Vision, Autonomous Vehicle, REMUS, BlueView, Rapid Prototyping, acoustic image modeling, sonar image analysis

15. NUMBER OF PAGES
69

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

ACOUSTIC IMAGE MODELS FOR NAVIGATION WITH FORWARD-LOOKING SONARS

Theodore Masek Jr.
Civilian, Naval Postgraduate School
B.S., California State University Monterey Bay, 2004

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SOFTWARE ENGINEERING

from the

NAVAL POSTGRADUATE SCHOOL
December 2008

Author: Theodore Masek Jr.

Approved by: Mathias Kölsch
Thesis Advisor

Kevin Squire
Second Reader

Peter Denning
Chairman, Department of Computer Science

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Cost and miniaturization of autonomous unmanned vehicles (AUV) drive component reuse and better sensor data analysis. One such component is the forward looking sonar (FLS) which can be used for obstacle avoidance and to extract vehicle state information. However, autonomous feature extraction of images from the FLS is difficult due to the noise inherent in the sensor and the sensor's susceptibility to interference from other acoustic devices.

This thesis investigated techniques to detect and classify common acoustic noise artifacts and common objects in a single frame. Other techniques require three or more frames to filter objects from other noise sources. A combination of probabilistic and template-based models were used to successfully detect and classify acoustic noise and objects. One common noise source is the micro modem which was detected 100% of the time with 1% false positives. Objects such as the ocean floor were correctly classified more than 93% of the time in most sites.

Due to the short development time frame, the software was developed with a two-stage approach. First, a high level scripting language was used for rapid prototyping of different classification techniques. In order to meet the time-constrained requirements of the target software, the classification algorithms were encapsulated as C++ classes in an object oriented design once the desired techniques were identified.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1

A. EQUIPMENT 3

1. REMUS AUV 3

2. Sensors 3

3. BlueView Forward Looking Sonar 6

4. Secondary Controller 7

B. SOFTWARE DEVELOPMENT PROCESS 8

C. SOFTWARE DESIGN 10

D. SONAR FEATURE EXTRACTION 11

II. ACOUSTIC IMAGE MODELS 13

A. ACOUSTIC NOISE BACKGROUND MODEL 13

B. ACOUSTIC INTERFERENCE NOISE MODELS 14

1. Acoustic Pulse Noise 16

2. Modem 17

C. GROUND MODEL 20

III. EXPERIMENTS 23

A. ACOUSTIC IMAGE BACKGROUND MODELS 24

1. Acoustic Pulse Noise 26

2. Modem Noise 27

B. GROUND 28

IV. RESULTS 29

A. ACOUSTIC IMAGE BACKGROUND MODEL 29

1. Acoustic Pulse Noise 30

2. Modem Noise 34

B. GROUND 36

V. DISCUSSION **45**
A. ACOUSTIC IMAGE MODEL 45
B. SOFTWARE DEVELOPMENT 47
VI. CONCLUSIONS **49**
LIST OF REFERENCES **51**
INITIAL DISTRIBUTION LIST **53**

LIST OF FIGURES

1. Left: Course with AUV flying above the ground on collision course with rock obstacle. Right: Actual forward looking sonar image overlaid on the schematic scene. 2
2. Naval Postgraduate School REMUS 100 AUV with BlueView FLS . . . 4
3. Different sensors REMUS uses for navigation in shallow water [Ref. 4]. 6
4. BlueView 450X with nose cone removed to show the six transducer configuration. 8
5. Left: Sonar image of the ground and approaching rock wall in Cartesian space (462x333). Right: image is the same image in polar space (461x1024). 9
6. MOOS Module Class Diagram. This UML diagram was created after techniques were finalized in MATLAB. It served as a plan when developing C++ code. 10
7. Top Left: Pulse artifact from unknown source and ground in Cartesian space. Top Right: Same image in Polar Space. Bottom Left: Modem noise and ground in Cartesian space. Bottom Right: Same image in Polar Space. 15
8. Ten pulse templates tested. A)Pt_{sd} B)Pt_{pc} C)Pt_{lc} D)Pt_{ff} E)Pt_{av} F)Pb_{sd} G)Pb_{pc} H)Pb_{lc} I)Pb_{ff} J)Pb_{av} 18
9. Common pattern found in all images with modem noise 19
10. Five modem templates tested. A)M_{pc} B)M_{ff} C)M_{sd} D)M_{lc} E)M_{av} . . . 19
11. Five ground model templates tested. A)G_{sd} B)G_{pc} C)G_{lc} D)G_{ff} E)G_{av} 22
12. Sample Images from the four sites. Top Left: Fisherman Flats (FF) Top Right: Panama City (PC) Bottom Left: San Diego (SD) Bottom Right: Lovers Cover (LC). 24
13. Sample Cartesian image annotated with LabelMe. 25

14. Intensity histogram from background training set of 50 images. Extracted is the histogram from the 300th row with the corresponding Gaussian model fitted. 26
15. Graph of average RMS error of 1024 histograms for each of the three functions. 29
16. Left: Each of the 3 functions fit to the histogram. Right: Graph of the RMS error of each of the fittings. 30
17. Percentage of pixels that fit within five different background intensity models. 31
18. ROC curve of background classification. 31
19. ROC graph of top-pulse templates with NCC 32
20. ROC graph of top-pulse templates with SAD 33
21. ROC graph of bottom-pulse templates with NCC 34
22. ROC graph of bottom-pulse templates with SAD 35
23. ROC graph of pulse templates with statistic test 36
24. ROC graph of modem templates with NCC 37
25. ROC graph of modem templates with SAD 38
26. ROC graph of modem detection with statistic scores 39
27. Plot of weighted score of five template fit to Fisherman Flats dataset . 40
28. Plot of weighted score of five template fit to Lovers Cove dataset 41
29. Plot of weighted score of five template fit to San Diego dataset 42
30. Plot of weighted score of five template fit to Panama City dataset . . . 43
31. Left: Sonar Image with obstacle protruding from ground. Right: Processed Sonar Image. 43
32. Left: Sonar Imagewith faint obstacle approaching. Right: Processed Sonar. 44

LIST OF TABLES

I. REMUS 100 Technical Specifications [Ref. 11]. 5

II. 450X-R100 Technical Specifications (from 450X-R100 manual). 7

III. Characteristics of the four sampling sites, numbers are frame counts. . 23

IV. Images used in the pulse test set. 27

V. Images used in the modem test set. 28

VI. Ground template fit weighted score summary for Fisherman Flats. . . . 38

VII. Ground template fit weighted score summary for Lovers Cove. 39

VIII. Ground template fit weighted score summary for San Diego. 40

IX. Ground template fit weighted score summary for Panama City. 40

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank my advisor Mathias Kölsch for his time and patience as I worked through this. His guidance and insight were invaluable. I would also like to thank my second reader Kevin Squire who not only helped a lot with writing but also allowed me to have this opportunity.

Most of all I need to thank my family, Jessica, Avalon and Aura, who's love and support made this possible.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Autonomous underwater vehicles (AUVs) are becoming a common tool for scientific marine research, commercial sea exploration, and for the military. AUVs excel in missions that require speed and stealth, especially where a tethered Remotely Operated Vehicles (ROV) to a surface ship would be impractical, too expensive, or too dangerous.

A common surveying tool for AUVs is a side scan sonar. This is an active sonar that is capable of rapidly surveying large swaths of the sea floor. During the sonar survey, the vehicle needs to maintain a constant altitude above the ocean ground. Depending on the specific side scan sonar device the survey altitude varies. For example, a Hydroid REMUS AUV with a Marine Sonics 900 kHz side scan sonar requires an altitude of three meters, while the current sensor suite on board the REMUS 100 AUV is able to react to ground slope changes of no more than 45° [Ref. 10].

Operating at low altitude, the AUV is constantly at risk of colliding with obstacles on the ground or floating in the water, such as storm debris. Forward looking sonar (FLS) can be used as an effective tool for avoiding possible collisions. Fig. 1 shows what an underwater obstacle avoidance scenario would look like. Recently, our lab has made progress in on-board image processing of BlueView 450 FLS for use in AUV obstacle avoidance.

FLS is ideal for obstacle avoidance because of the large area ensonified in a single ping. A ping in the horizontal plane can detect objects greater than 90 meters ahead of the vehicle in a 95°swath. However, autonomously extracting the objects from the image is very difficult. First, the size and shape of the possible obstacles is unknown. Furthermore, background noise due to gain and backscatter can hide objects that have a faint return, and the sonar is susceptible to artificially introduced artifacts caused by other acoustic sources such as the micro modem.

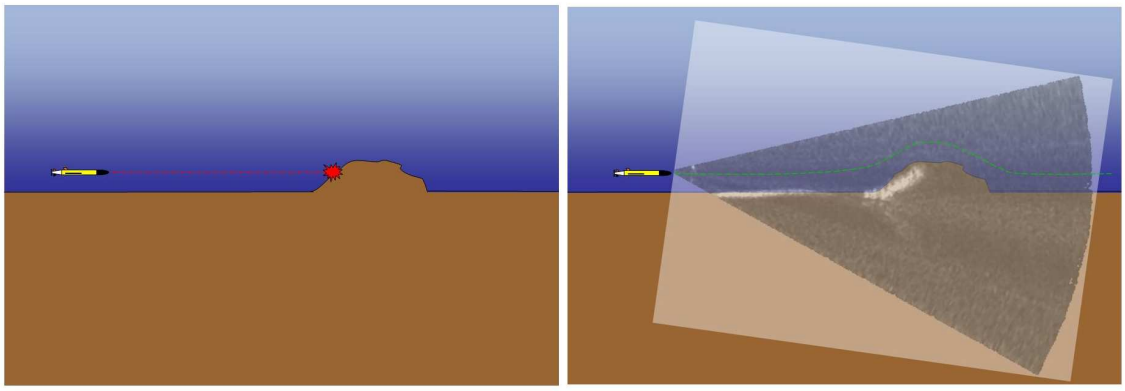

Figure 1. Left: Course with AUV flying above the ground on collision course with rock obstacle. Right: Actual forward looking sonar image overlaid on the schematic scene.

Our approach to detecting these unknown objects was to first model and identify common artifacts in the forward looking sonar image. There are three common artifacts in the image. The most prevalent is the background noise, followed by artifacts generated by the micro modem and a pulse-like artifact generated from an unknown source. Once these had been described we searched for common objects seen in the image. Currently the only object modeled is the ground. When operating at low altitudes the ground is present in almost every frame. Proper identification of the ground location and orientation has the benefit of being able to extract vehicle state information, such as pitch and altitude. With power limiting most AUV operations, the ability to save power by reducing sensors will increase mission times. Identifying the ground has the additional benefit of being able describe an obstacle's location to the vehicle and its altitude without needing any vehicle state information. After these steps are completed, we are able detect obstacles of unknown appearance with improved speed and accuracy.

The following models are described and applied to acoustic images of forward-looking sonars for detection and identification of noise, ground and obstacles. First, equipment-related work and prior efforts at acoustic image analysis and sonar-based

navigation are discussed. Next, the models are introduced, followed by a description of the experiments. The results and their discussion conclude this thesis.

A. EQUIPMENT

1. REMUS AUV

The AUV used for the experiments was the Remote Environmental Monitoring UnitS (REMUS) 100 Fig. 2. The REMUS vehicle was originally developed at Woods Hole Oceanographic institute, and is currently developed and sold through Hydroid Inc. The REMUS 100 are used in a variety of applications including hydrographic surveys, harbor security, environmental monitoring and mapping. One primary use for the NAVY is for mine counter measures (MCM).

The REMUS 100 is a small compact AUV that is 19 cm in diameter and weighs less then 37 kg. This enables the vehicle to be launched from any size boat without special deploying or recovery equipment. The vehicle can operate to a depth of 100 meters for up to 22 hours at speeds between 1.5-2.6 m/s. Refer to Table I for more details.

To navigate the REMUS 100 uses a combination of sensors Fig. 3. When operating on the surface, the REMUS 100 uses GPS to get a location fix. Once underwater, the REMUS measures its altitude and speed over the ground and speed through the water column with an Acoustic Doppler Current Profiler (ADCP)/Doppler Velocity Log (DVL). Along with a compass, the vehicle will navigate using dead reckoning. Dead reckoning alone is prone to position errors [Ref. 1]. To reduce these errors the vehicle will periodically update its location through a Long Base Line (LBL) fix. For LBL navigation the vehicle calculates its position by measuring the 2-way travel time of pings to acoustic beacons.

2. Sensors

The Naval Postgraduate School REMUS AUV is equipped with a standard set of passive sensors, including accelerometers and gyroscopes, to measure vehicle

Figure 2. Naval Postgraduate School REMUS 100 AUV with BlueView FLS

Table I. REMUS 100 Technical Specifications [Ref. 11].

Maximum Diameter	19 cm
Maximum Length	160 cm
Weight In Air	37 kg
Trim Weight	1 kg
Maximum Operating Depth	100 m
Energy	1 kw-hr Lithium Ion Battery
Endurance	22 hrs at optimum speed of 1.5 m/s 8 hrs at 2.6 m/s
Propulsion	Direct drive DC brushless motor to open 3-bladed propeller
Velocity Range	Up to 2.6 m/s
Control	2 coupled yaw and pitch fins
On / Off	Magnetic switch
External Hook-Up	2-pin combined Ethernet, vehicle power and charging 4-pin serial connector
Navigation	Long baseline (LBL) Ultra short baseline (USBL) and Doppler-assisted dead reckoning
Transponders	20 to 30 kHz operating frequency range
Tracking	Emergency transponder mission abort and in-mission tracking capabilites
Standard Sensors	ADCP/DVL Side Scan Sonar Long baseline (LBL), Ultra short baseline (USBL) Conductivity Temperature Pressure

state along with sensors to measure water conductivity, temperature, and depth. The REMUS carries acoustics sensors to calculate velocities, currents, and position. The Acoustic Doppler Current Profiler (ADCP)/Doppler Velocity Log (DVL) operates at 1200kHz and utilizes the Doppler Effect to measure water currents above and below the vehicle. The DVL is also used for altitude and velocity over ground estimation.

For mapping, the REMUS uses a Marine Sonics 900 kHz sidescan sonar. The Marine Sonics 900 has a range of 30 meters to each side of the vehicle when flying at a 3 meter altitude. The REMUS can send status messages back to the operator and

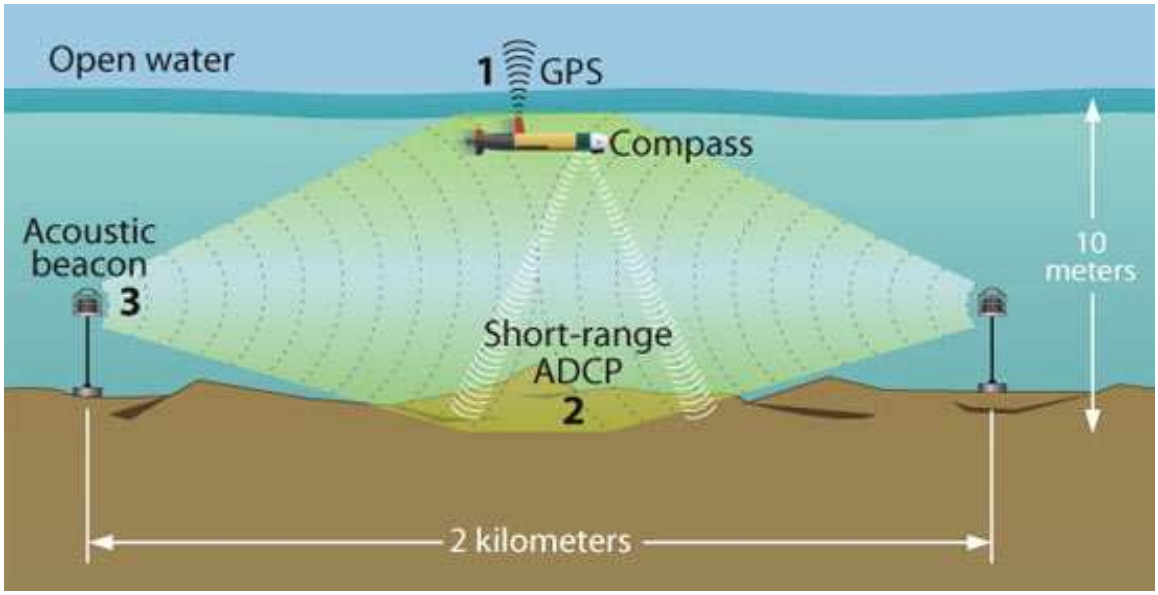

Figure 3. Different sensors REMUS uses for navigation in shallow water [Ref. 4].

receive start and stop commands through the acoustic Micro-modem. The acoustic Micro-modem operates at 25kHz and can transmit data between 80-5400 bps. The Micro-modem can also be used by the operator to get a range to the vehicle.

3. BlueView Forward Looking Sonar

The primary sensor used for this thesis was the BlueView 450X-R100 Forward-looking sonar (FLS) Fig. 4 . The BlueView 450X is a low power high-resolution blazed array sonar, comprised of a series of six small sonar transducers. Four are mounted in the horizontal plane and provide a 95°field of view. The remaining two transducers are vertically mounted and combine to provide a 45°field of view. Refer to Table II for more details.

With transducers mounted in the vertical plane, a better measurement of an object's height can be made. A common technique is to measure the shadow the object creates. However, if the object is higher than the vehicle's altitude then there is no shadow in the sonar image. The vertical plane can then be used to calculate the object's altitude. For this reason, the vertical plane was used as the primary obstacle

Table II. 450X-R100 Technical Specifications (from 450X-R100 manual).

Sonar Characteristics	
Max Range	137m (450ft)
Update Rate	Up to 10Hz
Swath Width	95°Horizontal 46°Vertical
Beam Width	1°x 15°
Electrical	
Power	12-48 volts DC @ 20 watts
Communications	Ethernet
Mechanical	
Depth Rating	100m
Weight in Air	9.57kg
Weight in Water	.95kg (positive)
Length w/ Nose Cone	53.97cm
Width	19.05cm
Acoustic	
Operating Frequency	300 to 600kHz

detection plane.

The sonar return can be visualized in Cartesian space or in polar space, see Fig. 5. The size of the polar space image from the sonar is 461x1024 pixels high. The sonar also returns a 462x333 Cartesian space image.

4. Secondary Controller

The REMUS has an on-board proprietary autopilot that controls the vehicle to accomplish a predetermined mission. The autopilot contains an interface that allows a second on-board user computer to receive vehicle state and status information and to take control of the vehicle in a limited fashion. The second computer in the Naval Postgraduate School's REMUS 100 is referred to as the secondary controller and is located just behind the BluewView FLS.

The secondary controller is an Intel 1.9GHz PC-104 with 1GB of RAM running GNU\Linux with a 2.6.24 soft real time kernel. The secondary controller on the REMUS vehicle has three primary functions: operating and storing images from

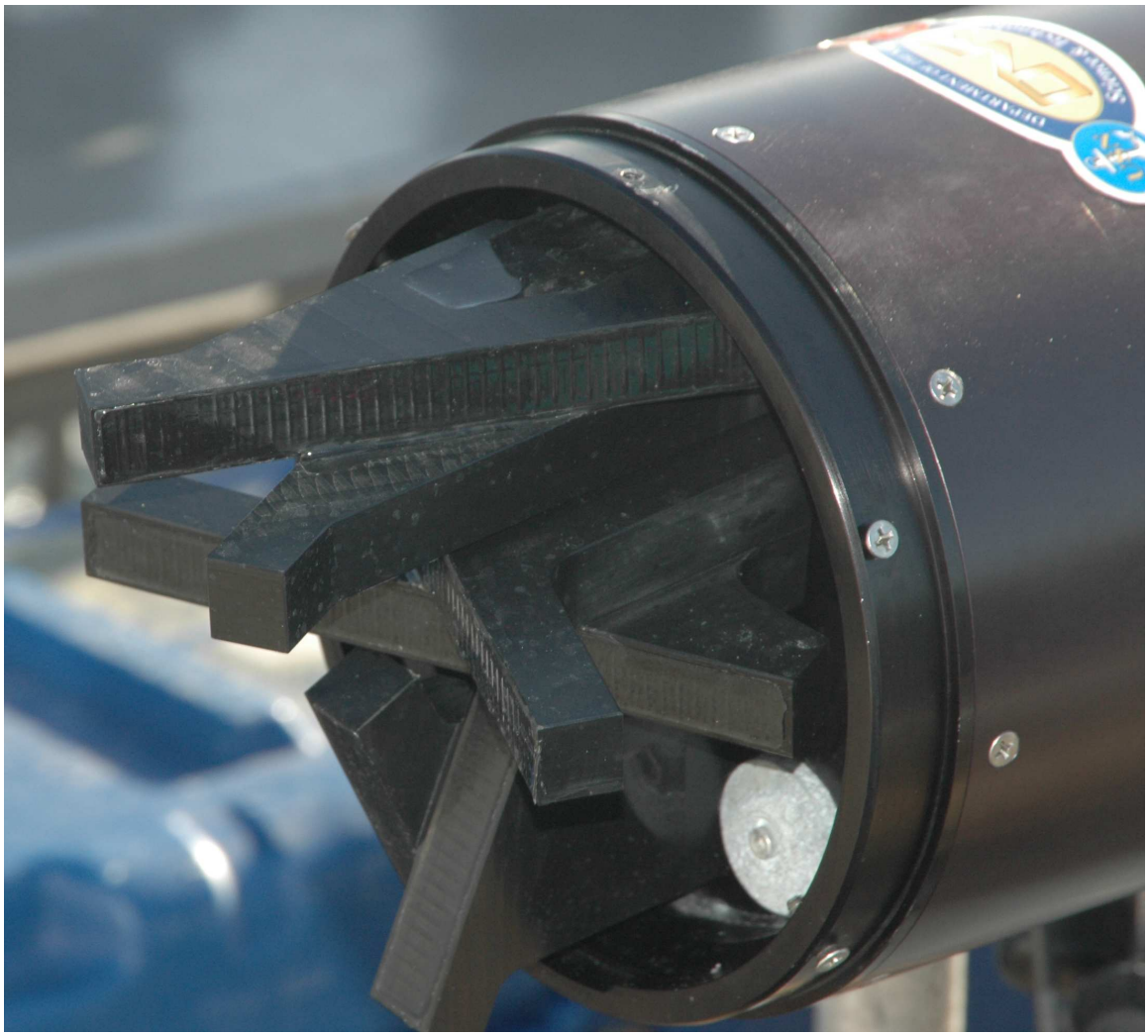

Figure 4. BlueView 450X with nose cone removed to show the six transducer configuration.

the FLS, monitoring and logging REMUS state information, and sending commands to the REMUS autopilot. Control of the sonar and REMUS along path planning and other tasks are comprised of smaller programs that are linked through Mission Oriented Operating Suite (MOOS) publish and subscribe architecture [Ref. 2].

B. SOFTWARE DEVELOPMENT PROCESS

Because of the short development time frame a two-step approach was taken to develop the software. First, the MATLAB scripting language was used for rapid

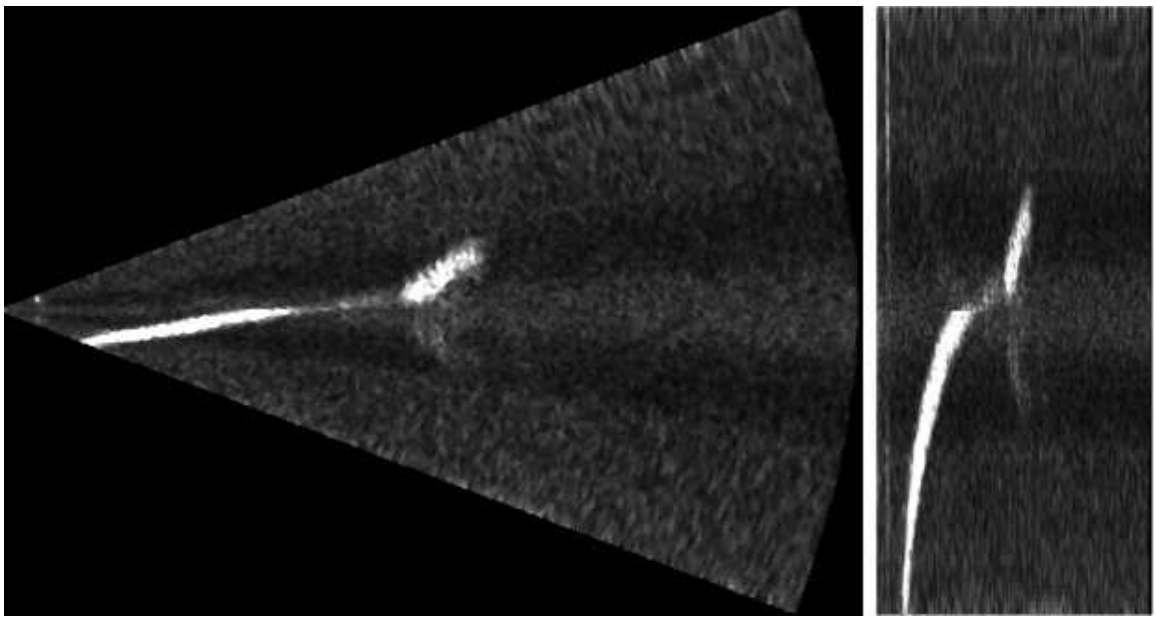

Figure 5. Left: Sonar image of the ground and approaching rock wall in Cartesian space (462x333). Right: image is the same image in polar space (461x1024).

prototyping of different models and techniques. Since MATLAB is a higher level scripting language, issues with memory management were alleviated. MATLAB also contains common linear algebra and image processing functions so different techniques were evaluated with much less effort. MATLAB allowed for quick generation of plots so results could be quickly evaluated in multiple ways. However, the ease of programming came at the cost of slow runtime speeds. Testing all four models could take up to a minute per image which is much longer than the 1.25 second refresh rate of the sonar.

The second step involved the porting of techniques developed in MATLAB into C++. This was accomplished by first creating a UML class diagram of the C++ class structure Fig. 6. The classification algorithms identified in Stage 1 were encapsulated as C++ classes in order for the target code to meet the timing constraints of the embedded system.

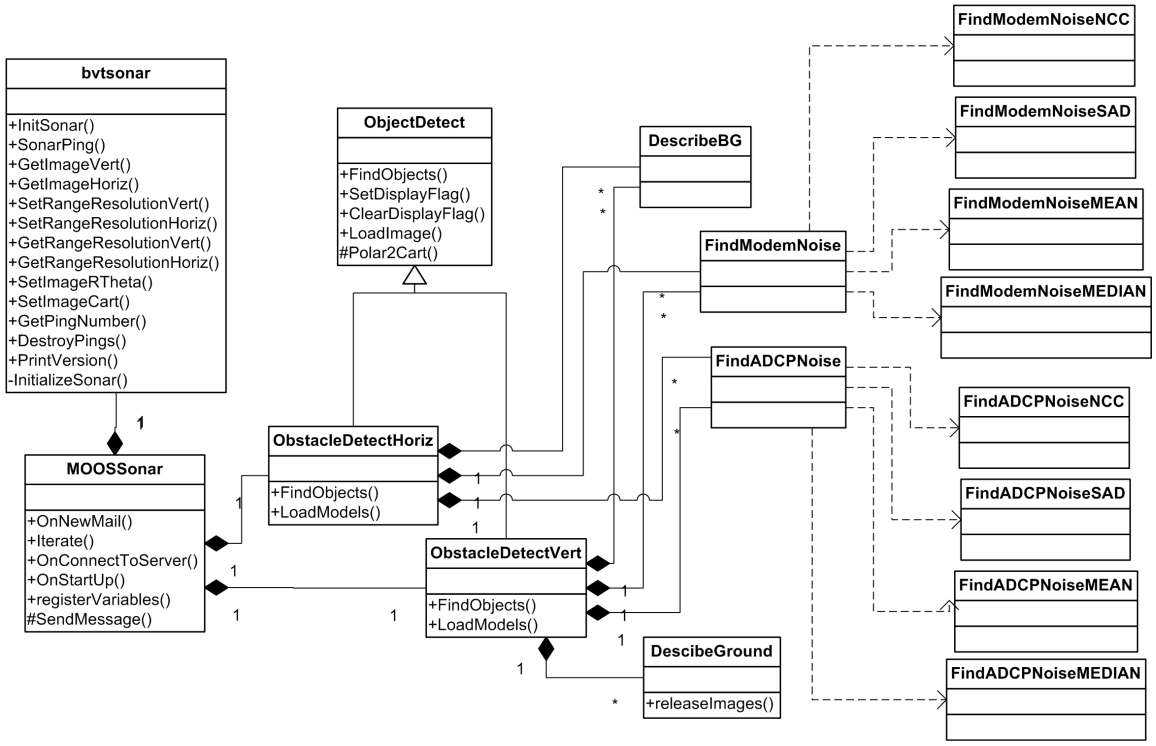

Figure 6. MOOS Module Class Diagram. This UML diagram was created after techniques were finalized in MATLAB. It served as a plan when developing C++ code.

C. SOFTWARE DESIGN

The MOOSSonar class is a subclass of the CMOOSApp class, refer to Fig. 6. The CMOOSApp class is an abstract class that describes the interface with the MOOS architecture. The MOOSSonar object instantiates bvtsonar to interface with the BlueView sonar, ObstacleDetectVert to extract objects from the vertical sonar images, and ObstacleDetectHoriz to extract objects from the horizontal image. ObstacleDetectVert and ObstacleDetectHoriz inherit common acoustic image processing methods from the ObjectDetect super class. The ObjectDetect class contains the functions to describe the background, detect and describe pulse noise and detect and describe modem noise. The models can be easily changed through a configuration file.

D. SONAR FEATURE EXTRACTION

Extracting features from side scan sonar images is a problematic process. A common approach to feature extraction is to locate and classify an object by locating its shadow [Ref. 16] [Ref. 17] [Ref. 13]. Using shadows from side scan images has two distinct advantages. First, the shadow regions are more invariant to changes in sonar conditions than the highlighted region [Ref. 18]. Second, shadows cast by man-made objects have a more regular geometric shape than shadows cast from natural objects [Ref. 13] [Ref. 19]. Some techniques used to extract shadows are snake models [Ref. 18] and template matching [Ref. 5]. Using shadows to extract features in side scan sonars has been applied to forward looking sonars as well [Ref. 7] [Ref. 22]. However, in our application, the use of shadows to detect possible obstacles has some limitations. The primary limitation is due to the mounting of our sonar (5 down tilt) and the low altitude the REMUS flies at, since the ground is only ensonified 30 meters ahead of the vehicle. When the ground beyond the object is not ensonified there is no shadow. Therefore only objects within a relatively short distance to the vehicle will create a shadow. Another limitation is that the object must be shorter than the vehicle's current altitude to create a shadow in order to calculate height. With objects taller than the vehicle's altitude the ground behind the object will never be ensonified, so there will never be a top to the shadow. Knowing the altitude of an object helps in determining the best way to avoid it.

Forward looking sonars can return multiple overlapping frames. Another common technique for feature extraction is to compare the movement of possible objects between the frames and filter them based on some predetermined parameters [Ref. 15] [Ref. 8] [Ref. 6] [Ref. 10]. While using multiple frames to filter obstacles has been successfully applied there are the following limitations. First, the technique requires the object to be imaged in multiple frames. The vehicle travels at 2.5 meters/second and the refresh rate for the 450X-R100 at a range of 90 meters is 1.25 seconds. If three frames are needed to classify an object the vehicle will travel roughly

6 meters before detecting it. This is a reasonable amount of time if the object is first seen 30-50 meters out, but this might be an issue when the vehicle is turning or pitching where new objects will be seen with much less distance to react. This assumption is valid in most cases, but violated in a shallow, high-traffic area such as a harbor where quick, reactive obstacle avoidance is needed.

Not only can FLS be used for obstacle detection and avoidance but vehicle state information and localization can be determined [Ref. 12] [Ref. 21] [Ref. 3]. In [Ref. 21]two-view homography was used with FLS images in the horizontal plane to estimate 3-D motion parameters. In [Ref. 21] EKF filters were used to extract useful features from the BlueView FLS to simultaneously localize the vehicle and create a map (SLAM).

II. ACOUSTIC IMAGE MODELS

The intensity of the echo that generates the acoustic image depends on many factors, including tiny particles in the ocean ("backscatter"), cross-talk between the transducers, other acoustic sources, the density of the object, and its size and distance.

Our approach to detecting various objects FLS images was to develop models of four of the common artifacts off-line, then apply these models to images in real-time when the vehicle is operating in the water. A blend of pixel-based and spatial (neighborhood, template-based) models were utilized depending on the object or effect. The four features commonly found in our data sets are background noise, noise caused by the micro-modem, a digital pulse caused by an unknown source, and the ground.

The four models were then applied to the original sonar image after smoothing with a 5x5 Gaussian filter. In order to extract features the image, noise and backscatter background was estimated, followed by a search for the presence of artifacts from other acoustic sources. Areas not well described by the first three techniques were fit with a ground model. Remaining pixels that were not well described were grouped by proximity to see if they were large enough to be pose a collision hazard. The complete model for the entire image is a mixture of source models P_i:

$$I(x, y) = \max(P_i(x, y))$$

A. ACOUSTIC NOISE BACKGROUND MODEL

The background is characterized by low intensity random signals. A uniform image-wide threshold is insufficient for describing the background because the average intensity of the background noise varies by angle across the field of view, a typical characteristic of this technology.

Using the polar plot and a set of images containing only the background, an intensity histogram was created for each angle. Visual observation suggested a Gaussian noise intensity model. Hence, a Gaussian was fit to the histogram, one per angle. The Gaussian would model the background intensity distribution for angle α, parameterized by its mean μ_α and variance σ_α^2. The score that a pixel was background was set to true (.98) if the intensity of the pixel fit within 3 standard deviations of the Gaussian models mean for that angle. The .98 score was determined from experimental results.

$$P_{bg}(\alpha, r) = \begin{cases} .98 \text{ if } |I_{\alpha,r} - \mu_\alpha| < 3\sigma_\alpha \\ 0 \text{ if } |I_{\alpha,r} - \mu_\alpha| > 3\sigma_\alpha \end{cases}$$

B. ACOUSTIC INTERFERENCE NOISE MODELS

There are two types of acoustic noise in the FLS images that are caused by a second acoustic source. The more common one is generated from the acoustic modem which occurs in about 7% of the frames. The acoustic modem fills about 80% of the image with small rectangles of varying intensities, see bottom of Fig. 7.

The other type of acoustic noise is a digital pulse that is caused by an unknown source. This artifact is less common, occurring in about 1% of the frames. The noise is less invasive, only affecting a small area of the image. See top of Fig. 7 for examples of the pulse artifact. Just as the background noise, both of these artifacts are easier to analyze in the polar space as they are found at specific angles and within a distance range.

A two-stage approach was used to detect and describe the acoustic noise. First, the image was quickly analyzed for noise presence. If detected, the appropriate model would "identify" pixels within an intensity range and location as the respective acoustic noise.

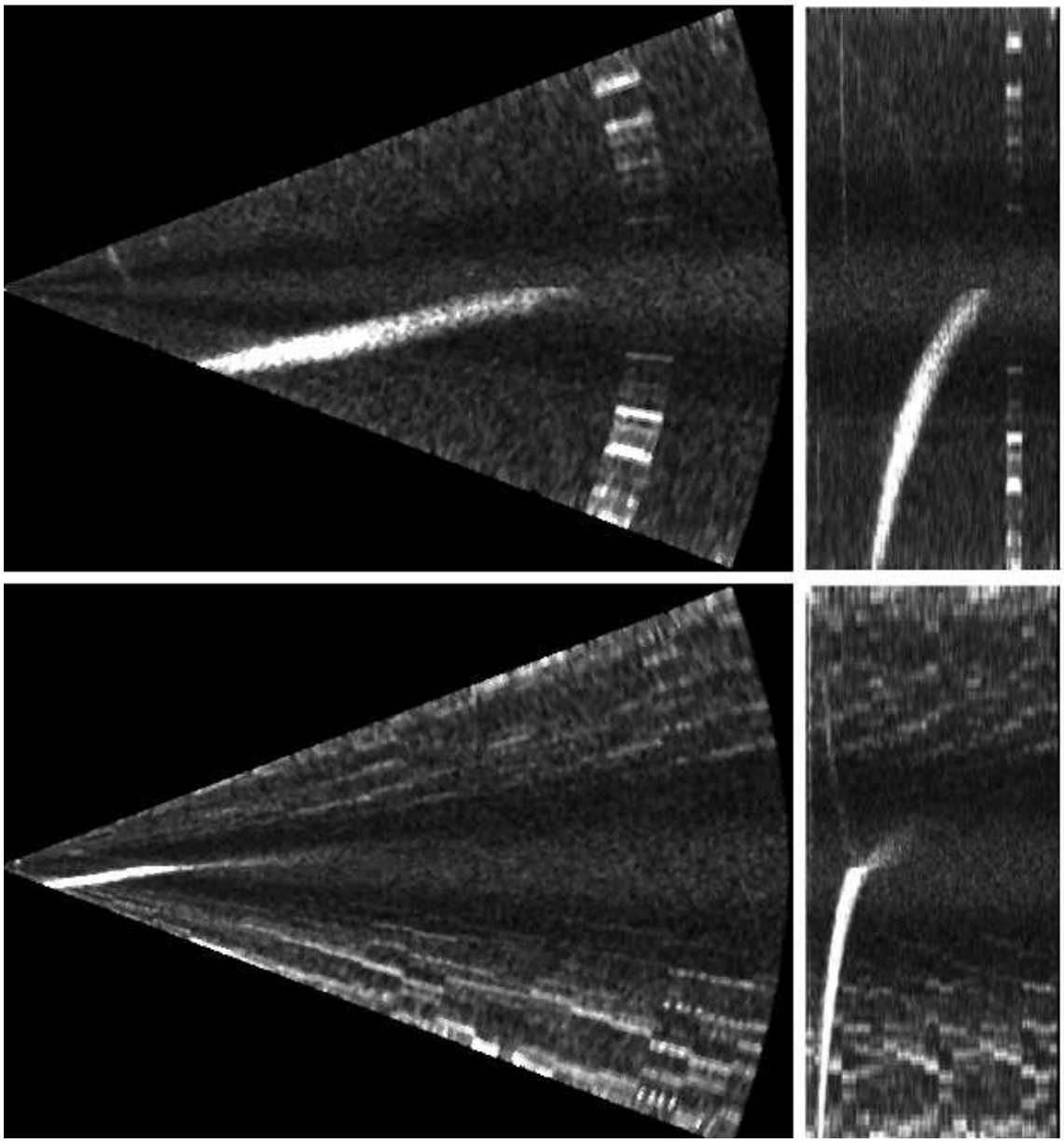

Figure 7. Top Left: Pulse artifact from unknown source and ground in Cartesian space. Top Right: Same image in Polar Space. Bottom Left: Modem noise and ground in Cartesian space. Bottom Right: Same image in Polar Space.

1. Acoustic Pulse Noise

The pulse noise has a digital pattern confined to a small band of approximately 75 pixels in width at a certain distance in the polar space. Under normal operations, the noise is slightly attenuated and is seen on the top ("top-pulse") and bottom ("bottom-pulse") third of the image. Since the full band is rarely seen, the search was limited to the top and bottom of the image. Because of the predictable size and shape of the search area, correlated template matching was used to detect the presence of pulse noise in the image. The template sizes for the top-pulse noise and bottom-pulse noise templates were determined separately as the average of all annotated pulse sub-images.

Four approaches were tested to detect images with pulse noise. The first two techniques applied templates of the noise to the image and scored the correlation using a sum of absolutes difference (SAD) and normalized cross-correlation (NCC) [Ref. 9].

$$SAD(x,y) = \sum_{i=0}^{rows} \sum_{j=0}^{cols} |(x+i, y+j) - (i,j)|$$

$$NCC(x,y) = \frac{\sum_{x,y} \left[f(x,y) - \bar{f}_{u,p}\right] \left[t(x-u, y-v) - \bar{t}\right]}{\left\{\sum_{\{x,y\}} \left[f(x,y) - \bar{f_{u,v}}\right]^2 \sum_{x,y} \left[t(x-u, y-v - \bar{t}\right]^2\right\}^{0.5}}$$

The lowest score from the SAD test and the highest score from the NCC test were used to detect the presence of pulse noise. The other two techniques calculate the mean and median of an area the size of the template on the top and bottom of the image. The 75x400 pixel area was shifted horizontally a row at a time until the mean and median were calculated for all the pixels across the top and bottom of the image. The highest score from each of the four sections (top mean, bottom mean, top median, bottom median) was used to detect the presence of pulse noise.

For the correlation test to detect top-pulse noise, five templates were tested. Four of the templates (Pt_{ff}, Pt_{lc}, Pt_{sd}, Pt_{pc}) were site-specific and created by averaging five sub-images containing only top-pulse noise from an individual site (Pt_{av}). The fifth template was created by averaging the four templates together, see Fig. 8

The above steps were repeated for creating five bottom-pulse templates (Pb_{ff}, Pb_{lc}, Pb_{sd}, Pb_{pc}, Pb_{av}).

The score from the detection test was used to classify images as containing pulse noise. Once an image was classified as having pulse noise the pixels within the area of the up-pulse and bottom-pulse template at the detected distance were set as being caused by pulse noise.

2. Modem

Four approaches were tested to detect images that contain modem noise. The first two techniques matched a template of an approximately 60 degree downward sloping "digital" pattern against parts of the image, about 3/4 down (see Fig. 9). The correlation was scored with SAD or with NCC. The second two techniques were the calculating of the mean and median intensities of the top 20% and bottom 20% of the image.

Five different templates were tested (M_{ff}, M_{lc}, M_{sd}, M_{pc}, M_{av}). The fist four were cropped directly from randomly sampled images from each site. The fifth was the average of the four templates together, see Fig. 10.

If modem noise was detected, pixels within an angle specific intensity range were marked as modem noise. This range was determined by finding the mean of the brightest pixels for each angle of a set of 20 annotated images only containing modem noise. The lower limit was three standard deviations above the mean of the background model for that angle.

For each angle the probability is set to 1 if intensity for the pixel (angle dist) is less than intensity threshold value, as in

$$P_{mn}(\alpha, r) = \begin{cases} 1 \text{ if } I(\alpha, r) > t_{bg} \text{ and } I(\alpha, r) < t_{mn} \\ 0 \text{ otherwise} \end{cases} .$$

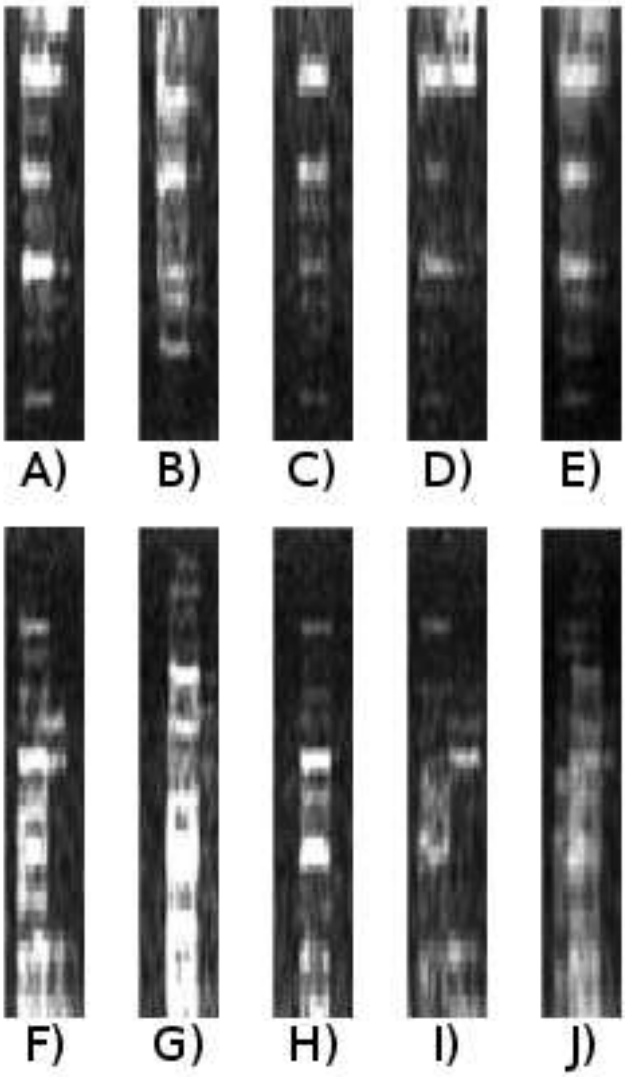

Figure 8. Ten pulse templates tested. A)Pt_{sd} B)Pt_{pc} C)Pt_{lc} D)Pt_{ff} E)Pt_{av} F)Pb_{sd} G)Pb_{pc} H)Pb_{lc} I)Pb_{ff} J)Pb_{av}

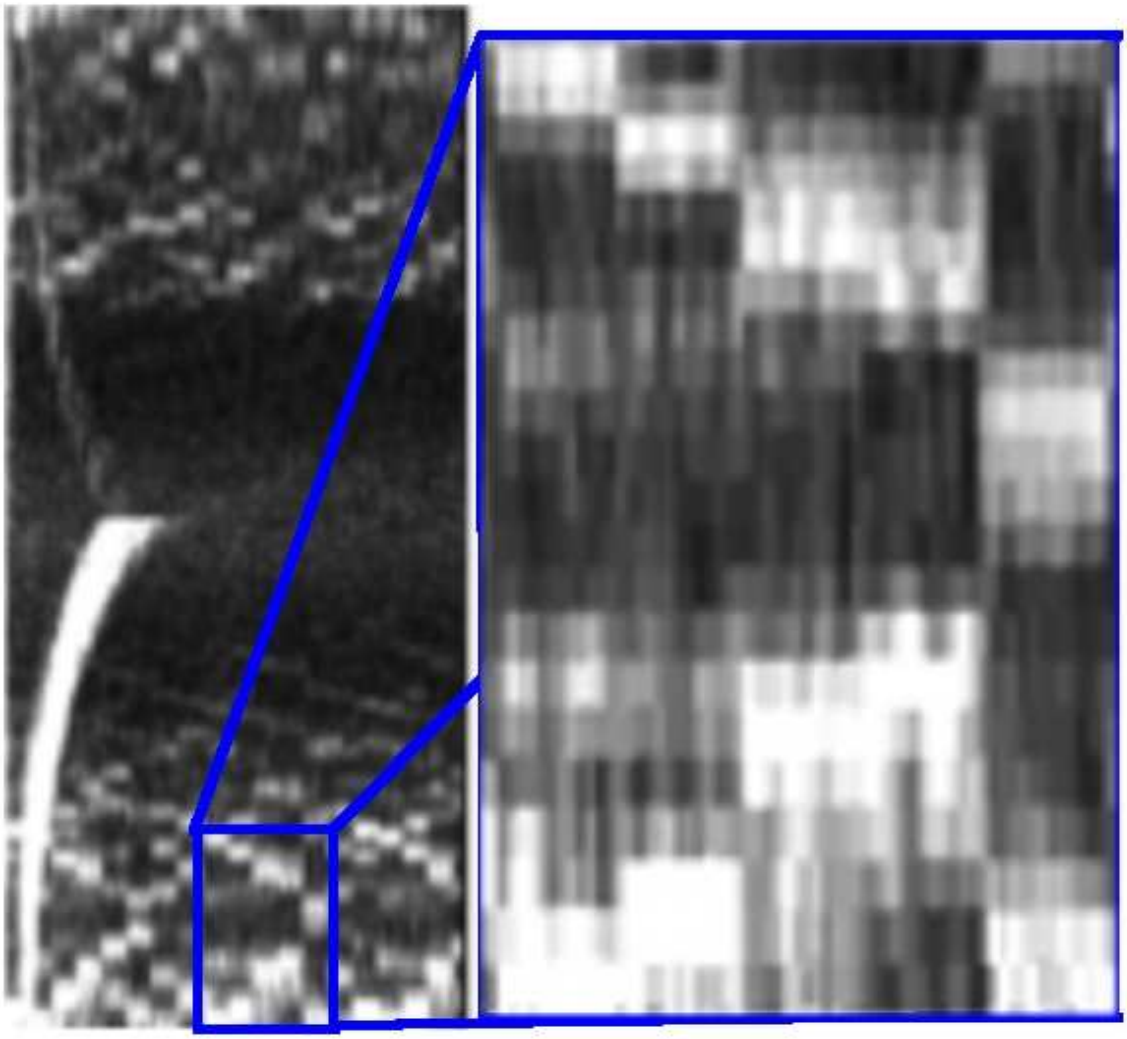

Figure 9. Common pattern found in all images with modem noise

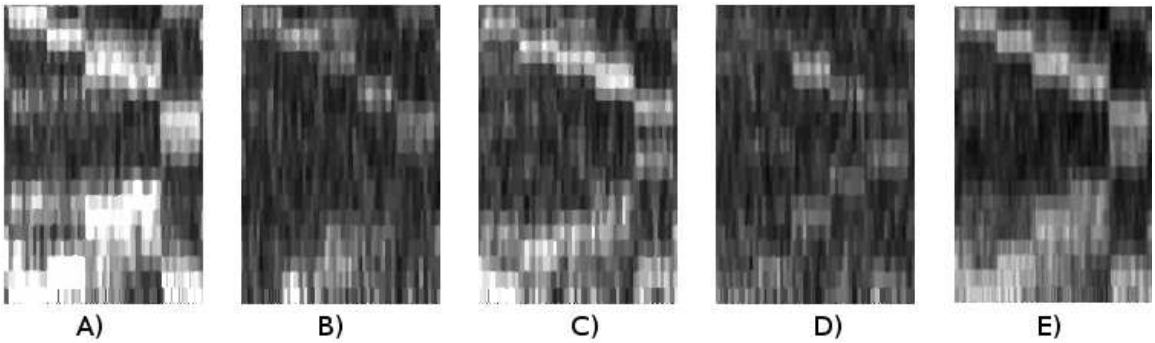

Figure 10. Five modem templates tested. A)M_{pc} B)M_{ff} C)M_{sd} D)M_{lc} E)M_{av}

C. GROUND MODEL

Unlike acoustic artifacts and the backscatter noise, the image of the ground changes position and rotation relative to the vehicle. The ground is also better represented in a Cartesian space image. The shape and intensity of the ground varies from different bottom types but is similar enough to describe with a common template. To fit the ground template over multiple environments, the template has five design parameters. It can be translated in x and y, rotated, and scaled by width and height independently. The template is scored on its fit using normalized cross-correlation. The template is fitted to the image using Nelder-Mead [Ref. 14] non-linear optimization method.

Five ground templates were created and tested (G_{ff}, G_{lc}, G_{sd}, G_{pc}, G_{av}). The first four were cropped manually from an image from each of the four sites. The image in each run was chosen to be a good representative to the given run. The fifth template was created by averaging forty ground images together, ten images from each site. Refer to Fig. 11 for the ground templates.

Before the ground model was fit to the image, initial points were needed. These start conditions were needed for the ground template because this technique is susceptible to issues with local minima. To locate initial conditions the previous noise description techniques were applied to the corresponding polar image. The polar image was then transformed to Cartesian space. Pixels not described with noise models were clustered by proximity. The center points of clusters larger than 300 pixels were used as the starting location for the ground template. The default template size was used for the initial scale. Six degrees was chosen as the initial rotation because the sonar has a five degree tilt down and the vehicle tends to operate with a one degree down tilt, making the ground appear at 6 degrees under the most common conditions.

The Nelder-Mead non-linear optimization method and NCC were used to fit the ground model to the original Cartesian image after it had been smoothed with a

5x5 Gaussian. The final location of the center of the template, final rotation, width scale, height scale and number of iterations were then recorded.

The important criteria to a good fit is that the top of the template ground and the annotated ground align. If the template is above the annotated ground, obstacles will appear lower and the vehicle might not classify it as a hazard. If the template is below the ground, then the vehicle and obstacle altitude appear high. Alignment of the top of the ground is also an important criteria because any information below the top surface of the ground is caused by the acoustic ping penetrating the medium and offers no information to vehicle or obstacle position. Therefore, as long as the template describes the entire ground it can be oversized in the horizontal and vertical scale.

To quantify the results a weighted score was used. To create the score a binary image was created of the fitted template and annotated ground. An intensity threshold of 59 was used because that was the average threshold for the background. The number of ground template pixels above the annotated ground ($Pixels_t$) and the number of annotated ground pixels above the ground template ($Pixels_a$) were summed together. These two criteria are only mutually exclusive when the template ground angle and annotated ground angle are equal. That score was then multiplied by the percent of annotated ground not covered by the ground template plus one (P_{g_not}).

$$e_{ground_fit} = (Pixels_t + Pixels_a) * (P_{g\_not} + 1)$$

This score measures how close the ground template and annotated ground are. If the ground template described the entire annotated ground, then the first two scores were multiplied by one. If the template and the annotated ground did not intersect at all, then the first two scores were summed and multiplied by two.

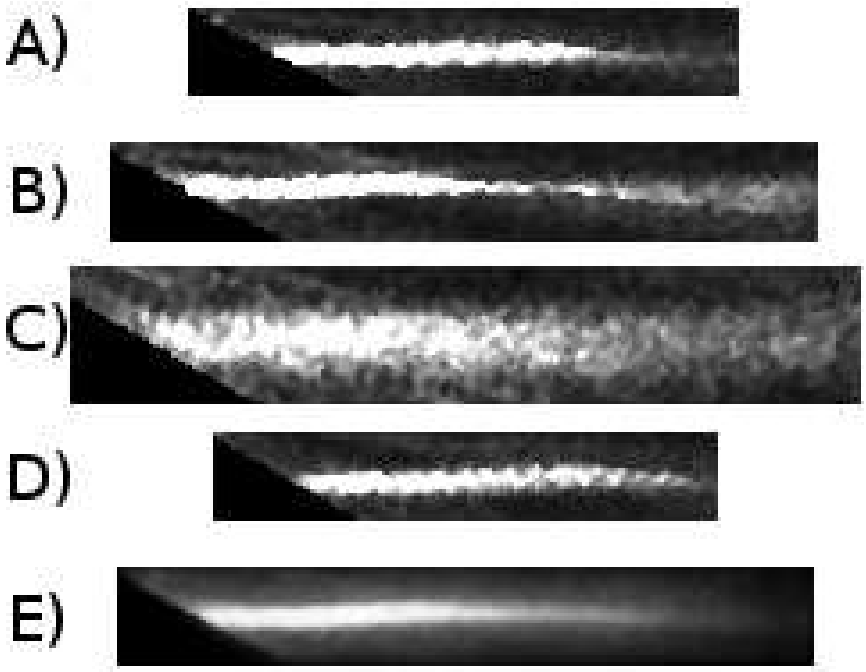

Figure 11. Five ground model templates tested. A)G_{sd} B)G_{pc} C)G_{lc} D)G_{ff} E)G_{av}

III. EXPERIMENTS

Data for the experiments were captured over the last year from four sites (see Table III), chosen due to varying factors such as bottom type and water temperature Fig. 12. Different sites were chosen to compare how well the models behaved in varying environments. The Fisherman Flats (FF) site in Moss Landing, CA was chosen because it contains objects the vehicle almost collides with. Silver Strand in San Diego, CA (SD) and Shell Island in Panama City, FL (PC) Florida have similar bottom types but with different environmental characteristics. The Lovers Cove (LC) site in Pacific grove, CA was used to test the models in very complex scenarios. Not only is this is a rocky shallow sight with both the ground and water surface in the frame but there is also a thick kelp forest present.

The data was logged to a file and later extracted on a PC without loss of quality compared to data obtained directly from the sonar head. The BlueView sonar returned a 16bit single channel image in a 463x333 pixel Cartesian space and a 16-bit image in polar space of 461x1024 pixels including the borders. The images were retrieved from the binary file using BlueView's API and OpenCV, then converted to 8-bit depth and saved as JPEGs. All the runs had the maximum sonar (ping) range set to 90 meters.

Table III. Characteristics of the four sampling sites, numbers are frame counts.

Site	Depth	Bottom Type	pings	modem	pulse
Fisherman Flats (FF)	15m	Sandy with 3m tall rock outcrop	1860	146	62
Lovers Cove (LC)	5m	Rocky with dense kelp forest	673	38	14
Shell Island (PC)	20m	Flat small grain sand	4384	332	142
Silver Strand (SD)	20m	Flat sand	5376	362	187

From these four sites roughly 300 images were annotated using the LabeMe tool [Ref. 20] by marking the silhouette of all objects that an experienced sonar image

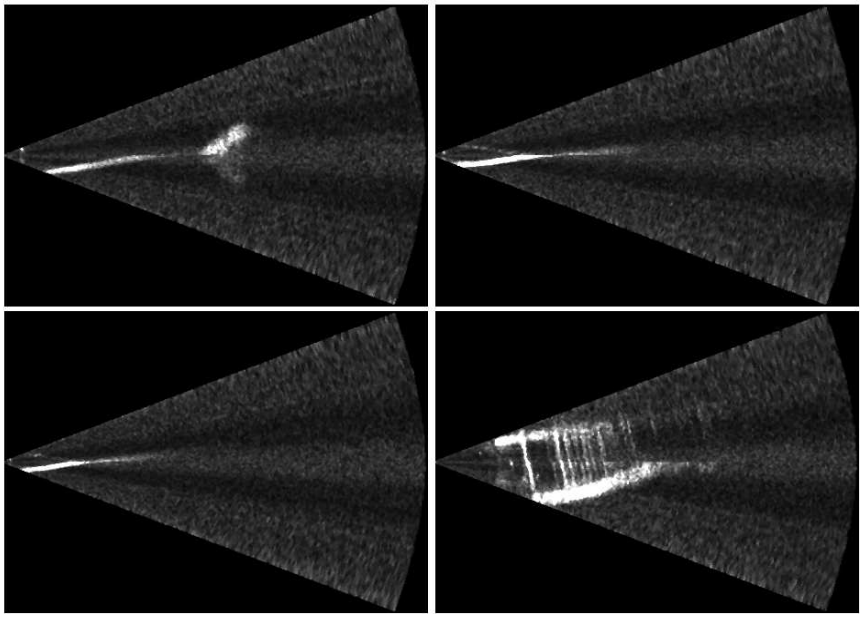

Figure 12. Sample Images from the four sites. Top Left: Fisherman Flats (FF) Top Right: Panama City (PC) Bottom Left: San Diego (SD) Bottom Right: Lovers Cover (LC).

analyst could find. Refer to Fig. 13 for a sample of an annotated image. Experiments were implemented in Matlab and run off-line.

A. ACOUSTIC IMAGE BACKGROUND MODELS

Five different background models were created and tested (Bg_1, Bg_{10}, Bg_{25}, Bg_{50}, Bg_{50a}). The first four models were created from randomly sampling 1, 10, 25 and 50 images from a set of 80 images (20 from each site) smoothed with a 5x5 Gaussian filter with all annotated areas set to 0. The fifth model was created from a dataset of 50 images that were taken from the sonar in the air (out of the water), which is thought to be a good measure for device-internal noise sources. A histogram was created for every angle in the polar images, resulting in 1024 histograms per model. Then a Gaussian curve was fit to each histogram $g(x) = a1 * exp(-((x - b1)/c1)^2)$. Refer to Fig. 14 for the set of histograms and Gaussian fit. Two other line-fitting functions were tested. An 8th degree polynomial $p(x) = \sum_{i=0}^{8} ai * x^{8-i}$ and an 8th

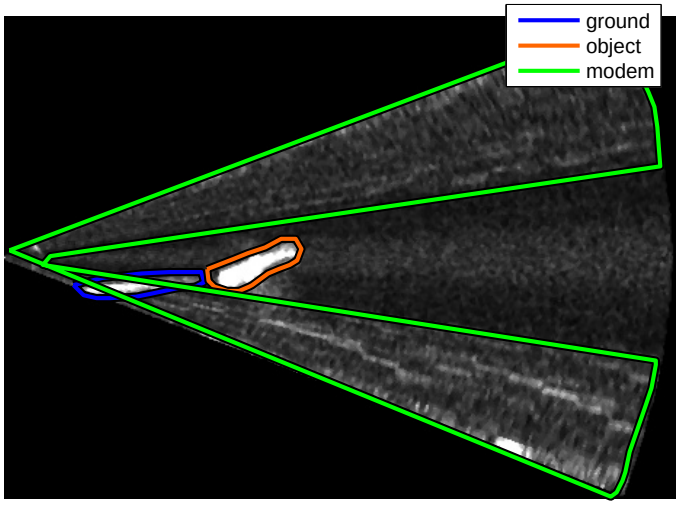

Figure 13. Sample Cartesian image annotated with LabelMe.

degree sum of sines $w(x) = \sum_{i=0}^{8} ai * sin(bi * x + ci)$ were also fit to the histogram. The Root Mean Square (RMS) error was calculated for each fit and averaged over all the histograms

The Gaussian background models were then tested against the 37 annotated images that had not been used to create the models. Again, to remove all artifacts that were not background noise, the images had the area of all tagged objects set to 0. An intensity of 0 only occurred in the border of the image. The quality of the model fit was determined by calculating the number of pixels with an intensity within three standard deviations of the model's mean and dividing the total number of pixels greater than zero for that row.

$$BgFit(\alpha) = \frac{\sum |I_{\alpha,r} - \mu_\alpha| < 3\sigma_\alpha}{\sum I_{\alpha,r} > 0}$$

The final background experiment was to test the model's ability to seperate background noise from other artifacts. The complete set of 87 images was used. First, a binary mask was created for each image. The binary mask set all pixels

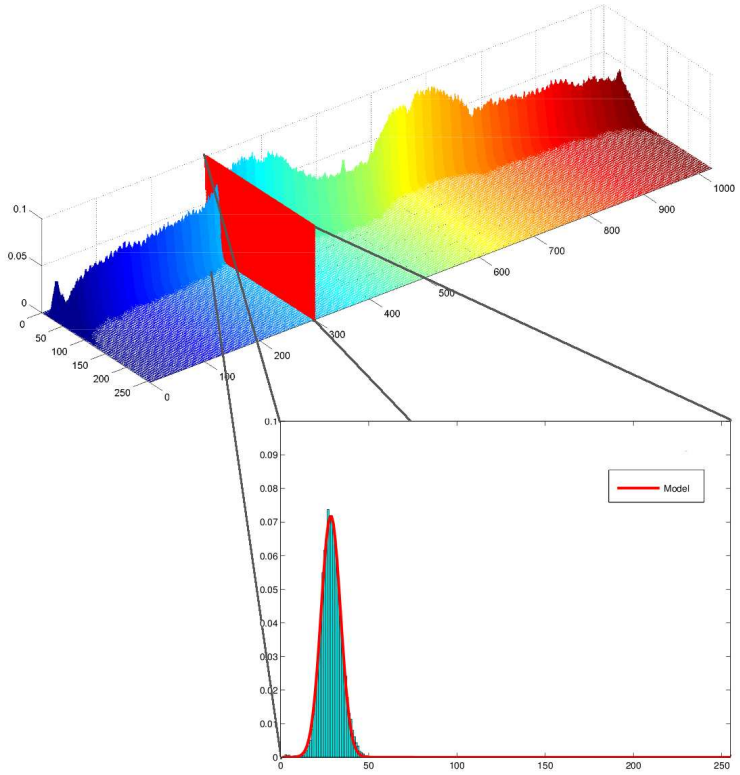

Figure 14. Intensity histogram from background training set of 50 images. Extracted is the histogram from the 300th row with the corresponding Gaussian model fitted.

within annotated areas to 1 and all other pixels (background) to 0. Next, a second mask was created by setting each pixel with an intensity within .01 deviations of the model's mean to 1. The dot product of the two masks resulted in a matrix where 1 represents a false positive. The not-background mask was subtracted from the background model mask to get the true positives.

1. Acoustic Pulse Noise

To detect pulse noise in an image 24 experiments were conducted on a 552 image test set. Refer to Table IV for what images comprised the test set. Twenty of the experiments were template matching. Ten of the experiments applied the five

top-pulse templates (Pt_{ff}, Pt_{lc}, Pt_{sd}, Pt_{pc}, (Pt_{av}) to the top of each of the 552 images, recording the highest NCC score and the lowest SAD score. The other ten experiments applied the five bottom-pulse templates (Pb_{ff}, Pb_{lc}, Pb_{sd}, Pb_{pc}, Pb_{av}) to the bottom of each image and the highest NCC score and the lowest SAD score were recorded. Before each experiment every image was smoothed with a 5x5 Gaussian filter.

Table IV. Images used in the pulse test set.

Site	pulse noise	no pulse noise
Fisherman Flats (FF)	62	62
Lovers Cove (LC)	14	14
Shell Island (PC)	100	100
Silver Strand (SD)	100	100
total:	276	276

The other technique consisted of four experiments calculating the mean and median of sections of the image. This technique was similar to the template matching. An area the size of the template was moved across the top and bottom of the image and the mean and median within the area were calculated. The highest score from each test was recorded.

2. Modem Noise

There were 12 experiments testing different ways to detect modem noise in a 676 image test set, refer to Table V for what images comprised the test set. Before each experiment every image was smoothed with a 5x5 Gaussian filter.

Two different techniques were tested to detect modem noise. The first technique was template matching and consisted of ten experiments. Five templates (M_{ff}, M_{lc}, M_{sd}, M_{pc}, M_{av}) were applied to the bottom 20% of each image and scored using NCC and SAD. The highest score from NCC and the lowest from SAD were recorded. The other two experiments calculated the mean and median of the top and bottom third of the image.

Table V. Images used in the modem test set.

Site	modem noise	no modem noise
Fisherman Flats (FF)	100	100
Lovers Cove (LC)	38	38
Shell Island (PC)	100	100
Silver Strand (SD)	100	100
total:	338	338

B. GROUND

The ground model dataset was comprised of 40 annotated Cartesian-space images, ten from each of the four sites. Each set from the different sites contained two images with modem noise and one image with pulse noise. Some of the images contained unclassified objects.

The five ground templates (G_{ff}, G_{lc}, G_{sd}, G_{pc}, G_{av}) were fit to the 40 test images. The weighted score was calculated for each of the 200 experiments.

$$e_{ground_{fit}} = (Pixels_t + Pixels_a) * (P_{g\_not} + 1)$$

IV. RESULTS

A. ACOUSTIC IMAGE BACKGROUND MODEL

Three functions were tested to see which best described the background intensity histogram. Each curve was fit to each of the 1024 histograms by minimizing the Root Mean Square (RMS) error. Refer to Fig. 16 for a sample of the curve fitting to one of the histograms. Refer to Fig. 15 of the average RMS error for each function over the 1024 histograms. The plot clearly shows that the Gaussian model has a much smaller RMS error then the other two models. The average RMS value was 0.000167 which is 0.42% of the max histogram count.

For each of the five histograms, models were applied to 30 test polar space images containing only background noise. The ratio of pixels that fit within the model was averaged for each angle over the 30 images. Refer to Fig. 17 for the mean and standard deviation of pixels described by the background model. All the Gaussian

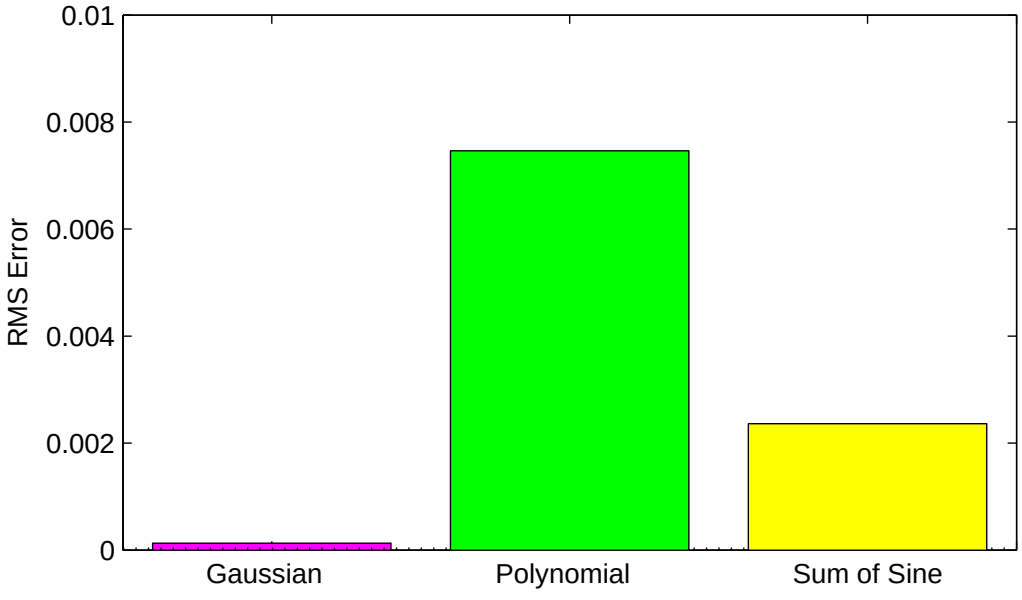

Figure 15. Graph of average RMS error of 1024 histograms for each of the three functions.

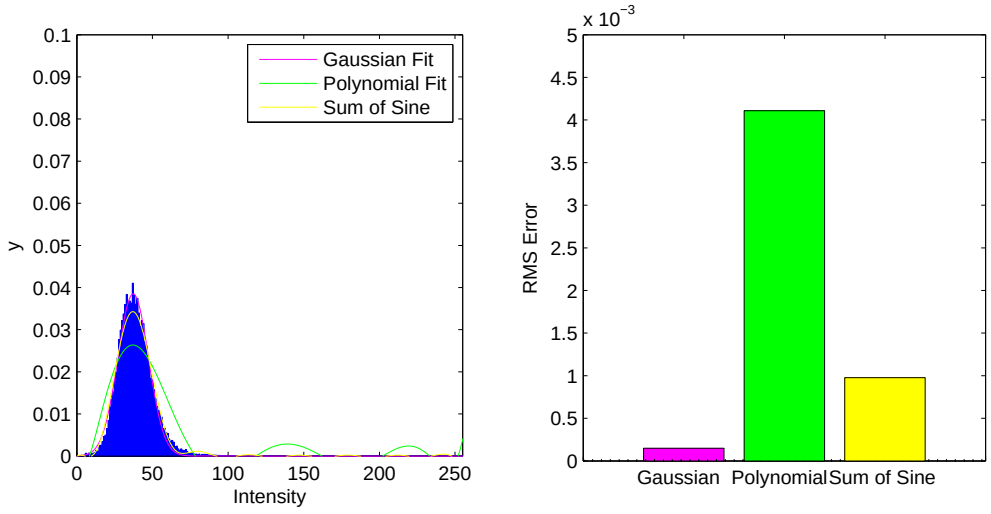

Figure 16. Left: Each of the 3 functions fit to the histogram. Right: Graph of the RMS error of each of the fittings.

model tests described more than 96.5% of the background pixels. While models created with 50 images performed better ,the average number of pixels correctly described was within half a percent. This means that training sets larger than ten images offer little benefit for creating background models. The models created from background noise in 50 images taken in the air performed the worst. This is possibly due to the models describing the background noise from the gain but not backscatter from particles in the water.

Fig. 18 show a ROC curve of the background classification. The ROC curve shows the percent of background classified to the percent of pixels that are not background misclassified, by changing the standard deviation of the model from .01 to 17. The curve shows that there is roughly 40% false positives to reach the 95% true positives. After analyzing the results the number of false positives appears to be high.

1. Acoustic Pulse Noise

Twenty-four experiments were conducted comparing two different techniques and two search areas to detect images with pulse noise. Figures 19,20,21,22,23 are the ROC curves that were created from each of the experiments. The ROC curves

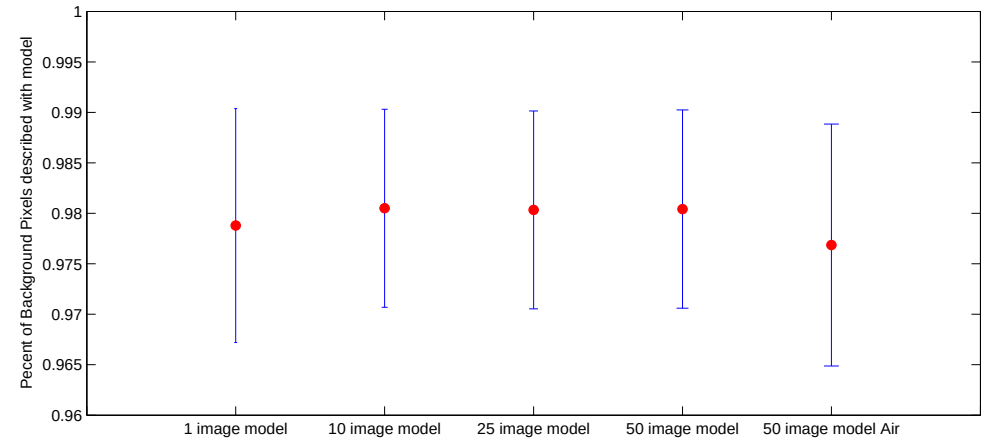

Figure 17. Percentage of pixels that fit within five different background intensity models.

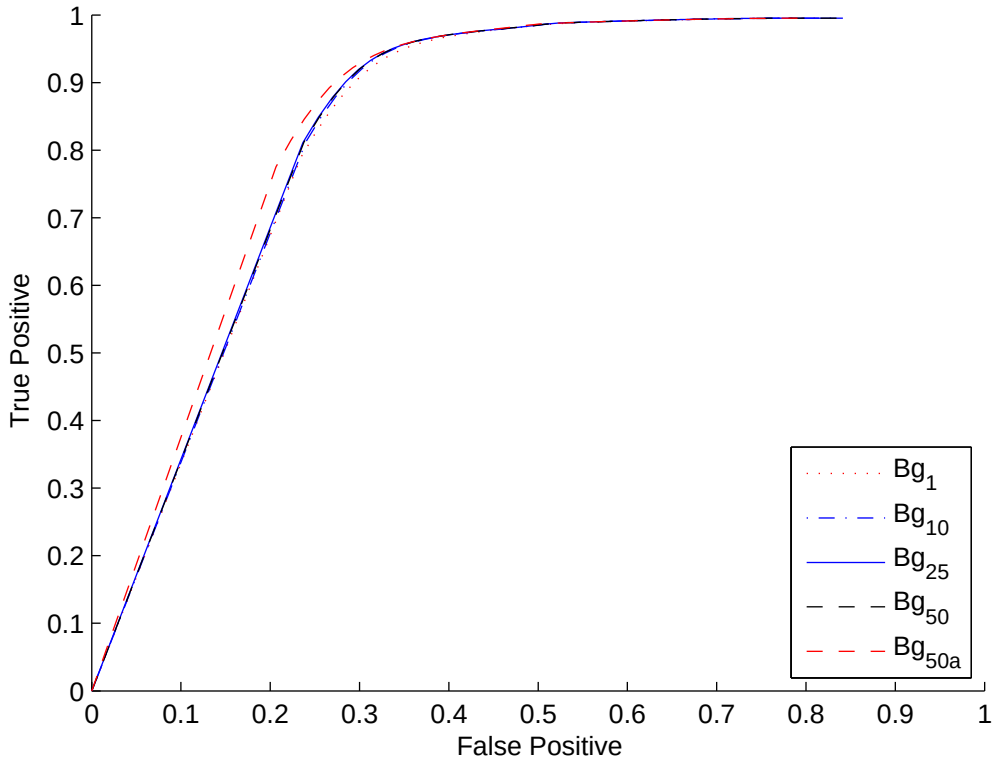

Figure 18. ROC curve of background classification.

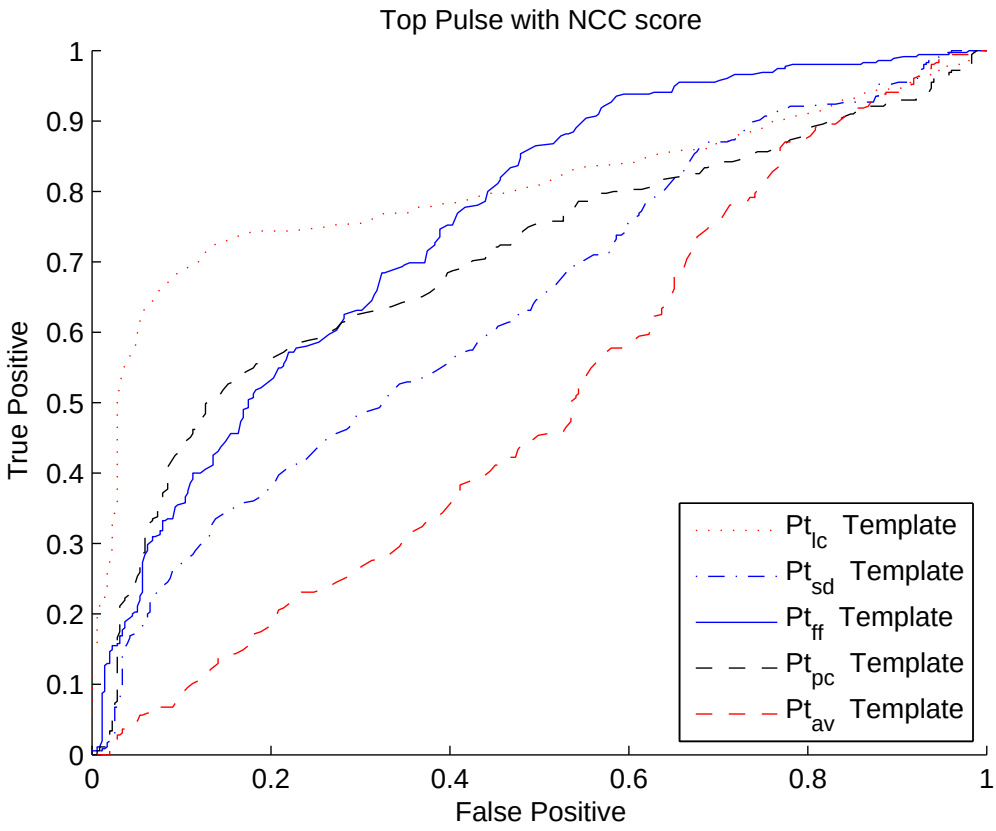

Figure 19. ROC graph of top-pulse templates with NCC

19,20,21,22 of template matching experiments show how many images were correctly described as having pulse noise to those that were incorrectly described as having pulse using the correlation score as the threshold. The other ROC curves in Fig. 23 show how many true positives to false positives were found at different intensity values from the mean and median experiments.

Three of the five templates detected pulse noise at the top of the image 90% of the time,(see Fig. 19). One surprising result was how poorly the Pt_{ff} template performed. This is probably due to the template being very bright at the top with relatively little data in the rest of the template.

When the template difference was calculated with SAD, the Pt_{sd} template performed very well Fig. 20. It detected images with top-pulse noise 92% of the

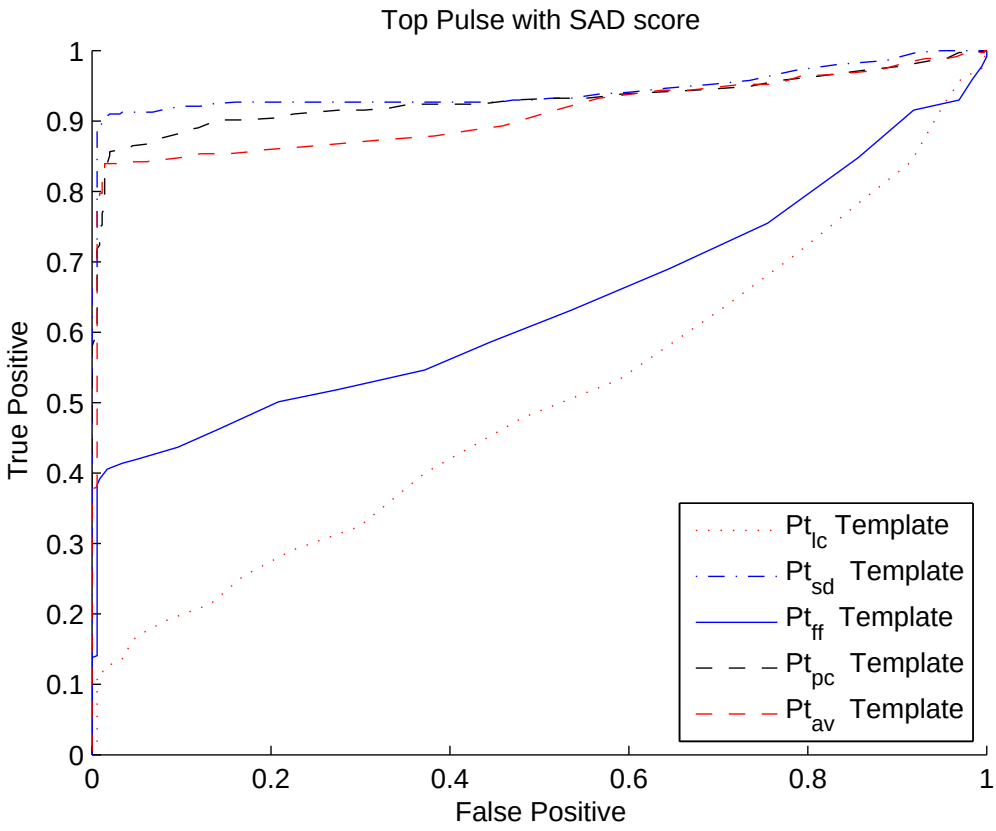

Figure 20. ROC graph of top-pulse templates with SAD

time. The Pt_{av} and Pt_{pc} templates were strong with both being able to detect noise greater than 83% with less than 3% false positives. Both thePt_{ff} and Pt_{lc} templates performed poorly, with Pt_{ff} showing less than 40% detection andPt_{lc} less than 15% detection before starting to get false positives.

None of the templates performed well with detecting bottom-pulse noise with NCC Fig. 21. The Pt_{lc} template was the best but it was only able to detect 15% of images with modem noise before it started detecting false positives.

Detection of bottom-pulse noise with SAD template matching is not a very robust way to detect pulse noise Fig. 22. While two templates (Pb_{pc} and Pb_{sd}) scored better than the others, they were still only able to detect 70% of the images with noise before detecting false positives. The other templates Pb_{ff}, Pb_{lc}, and Pb_{av} would only

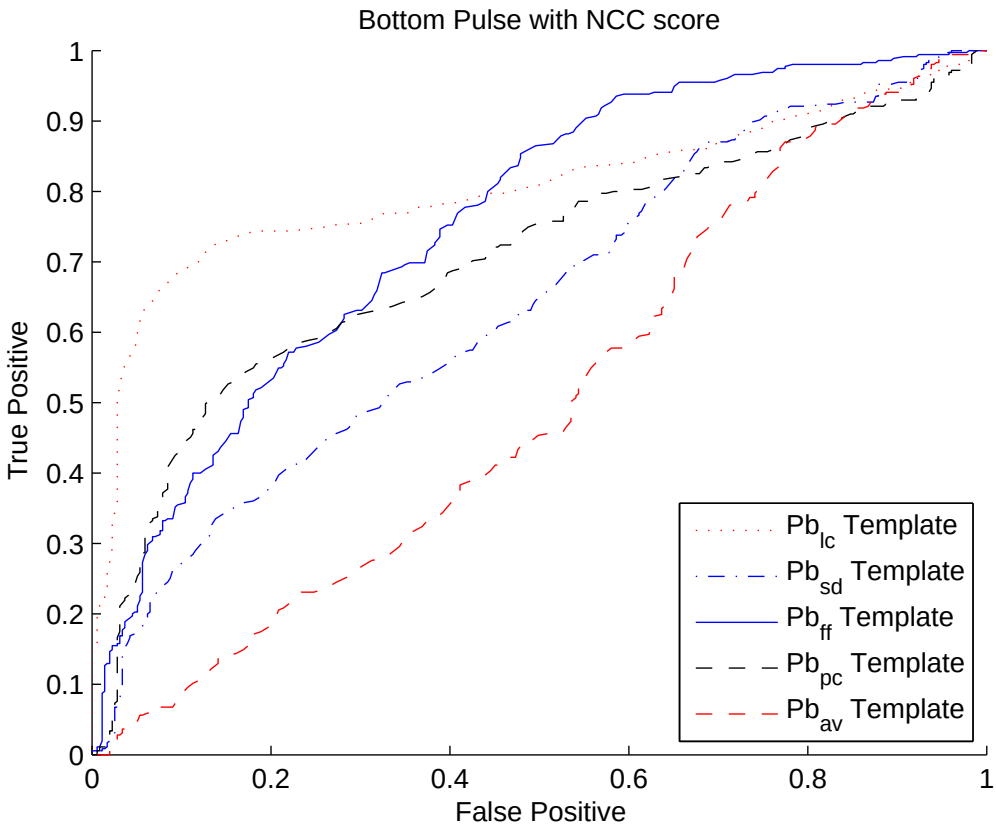

Figure 21. ROC graph of bottom-pulse templates with NCC

detect 11%, 28% and 40% respectively.

Using statistical tests is not a good way to detect images with pulse noise. Only after 5% false positives would the true positive detection rise.

2. Modem Noise

Twelve experiments were conducted comparing two different techniques to detect images with modem noise. Figures 24,25,26 are the ROC curves that were created from each of the experiments. Similar to the pulse detection test, the ROC curves of template-matching experiments show how many images were correctly described as having modem noise to those that were described as having modem noise and did not use the correlation score as the threshold. The statistical experiment ROC curves show how many true positives there are to false positives at different intensity values

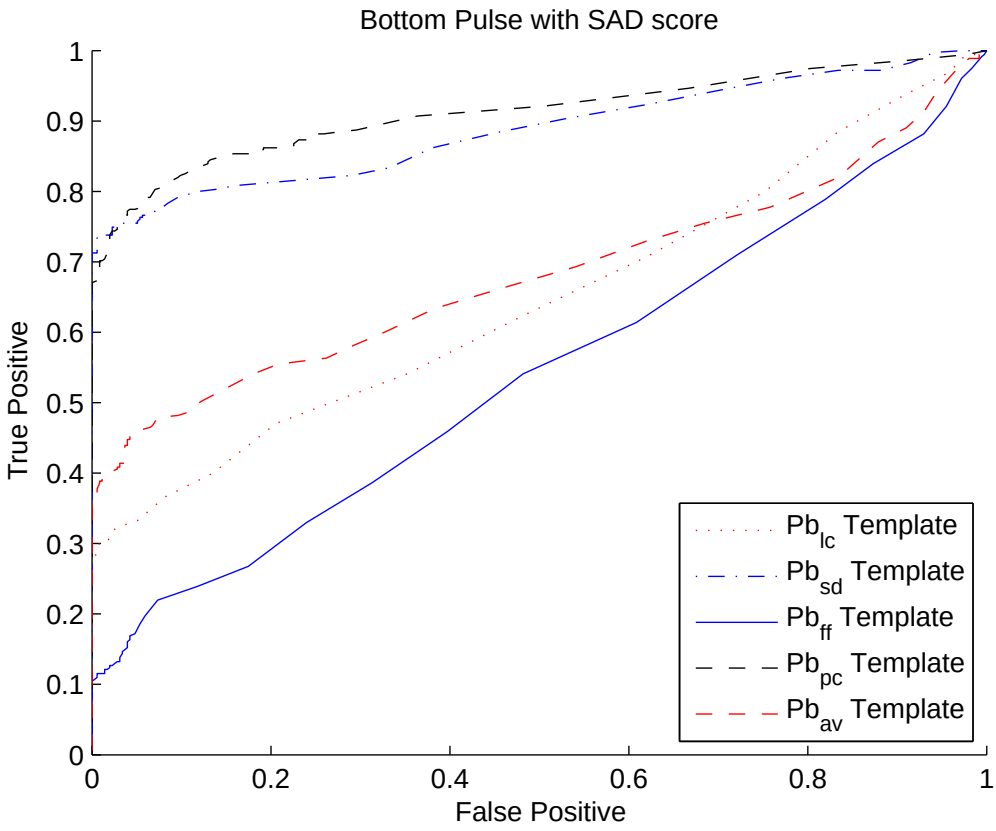

Figure 22. ROC graph of bottom-pulse templates with SAD

from the mean and median experiments.

All the modem templates scored with NCC performed well in detecting modem noise Fig. 24. The lowest-performing templates still detected 80% of the images with less than 3% false positives. The best-performing templates (M_{lc} and M_{av}) both detected 96% true positives before detecting false positives.

An interesting result was the ROC curve for modem templates scored with SAD Fig. 25. The M_{pc} template scored the best out of all the experiments with 100% detection and less than 3% false positives. The rest of the templates performed very poorly with very interesting curves. The unique flat curve is because of the templates brightness. The M_{lc} template was quite dim compared to theM_{pc}C template. When the scored with SAD the template was closer to background than the site with bright

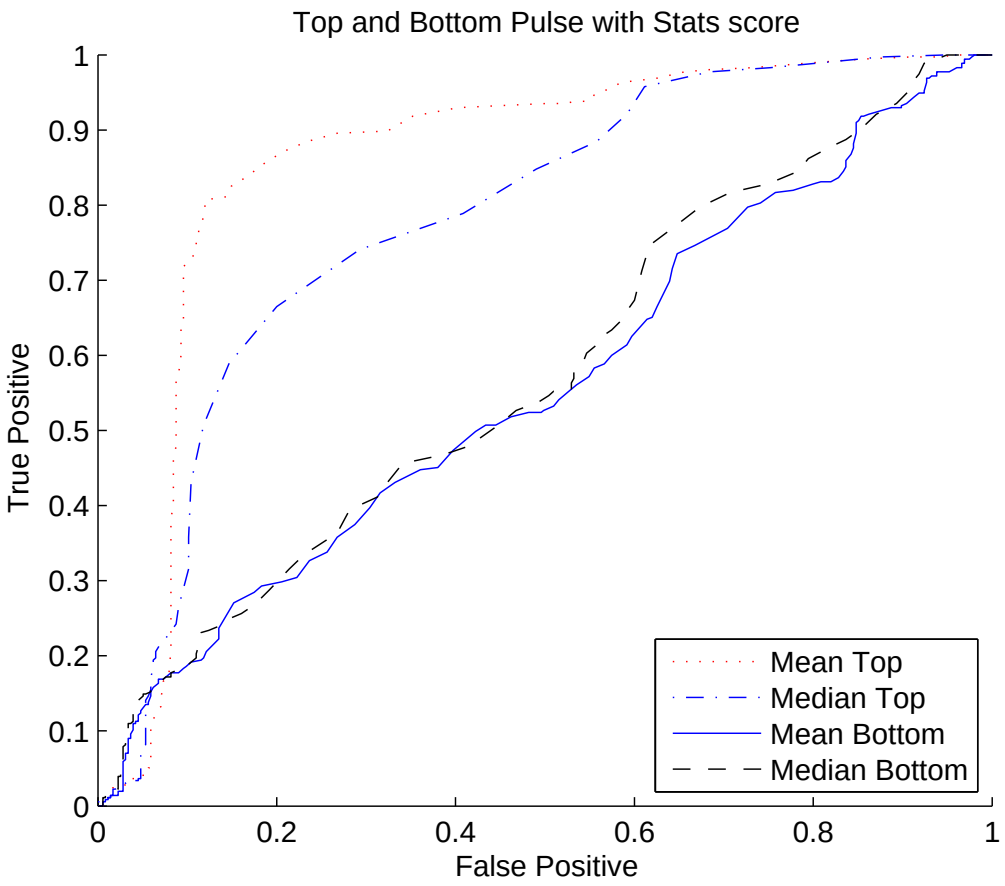

Figure 23. ROC graph of pulse templates with statistic test

modem noise. Therefore, the template worked well on one site, then the difference between the brightness of the background noise and the template was closer than the brightness of the template to the sites with bright modem noise.

The statistics test to detect modem noise performed much better than anticipated Fig. 26. The mean and the median detected 97% of images before detecting any false positives.

B. GROUND

The five ground templates (G_{ff}, G_{lc}, G_{sd}, G_{pc}, G_{av}) were fit to 40 test images representative of the different sites' characteristics. To quantify how well the ground

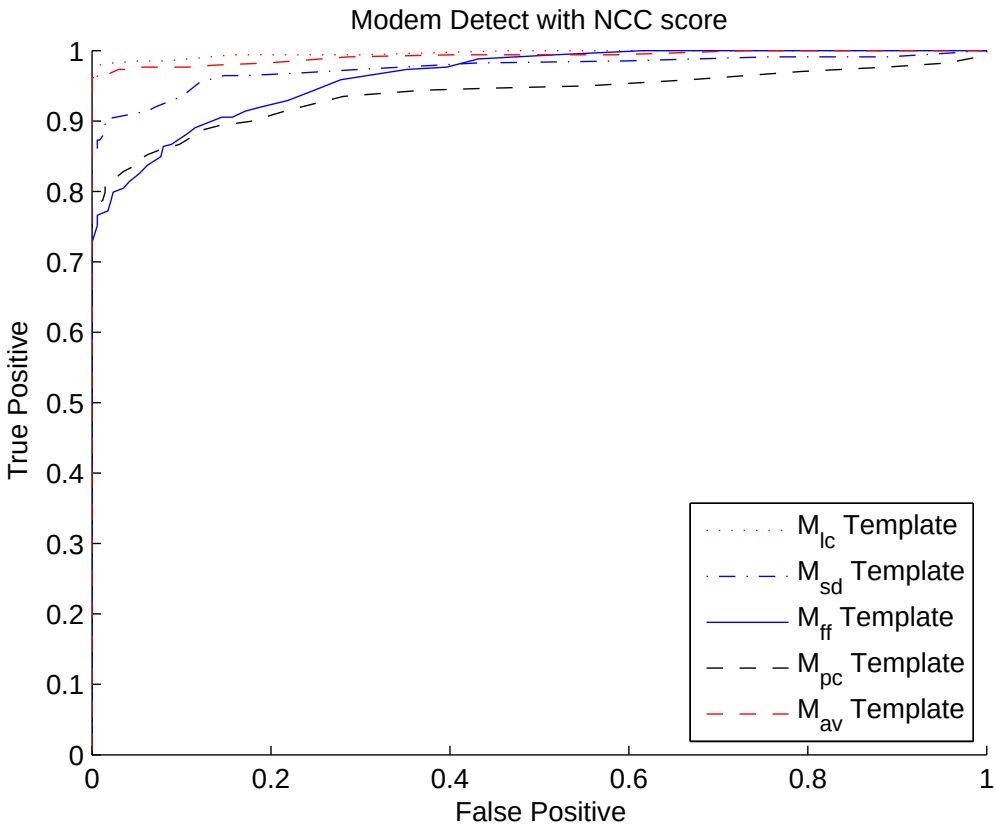

Figure 24. ROC graph of modem templates with NCC

template fit the annotated data the following weighted score was used $e_{ground_fit} = (Pixels_t + Pixels_a) * (P_{g\_not} + 1)$. After analyzing the results the following threshold values were used to classify how well the template fit. A score less than 200 was considered an excellent fit. To get a score of 200 the annotated ground must be completely described by the ground template and the tops of the ground can be shifted one row of pixels in the vertical direction. Scores less than 500 were considered a good fit, scores between 500 and 1000 were poor and any score greater than 1000 was an unacceptable fit. Four plots were created of the scores from each site, Fig. 27,28,29,30. Refer to Tables VI,VII,VIII,IX for a summary of how each template performed at the different sites. Fig. 31 and Fig. 32 are two sample plots of the completed mixture model.

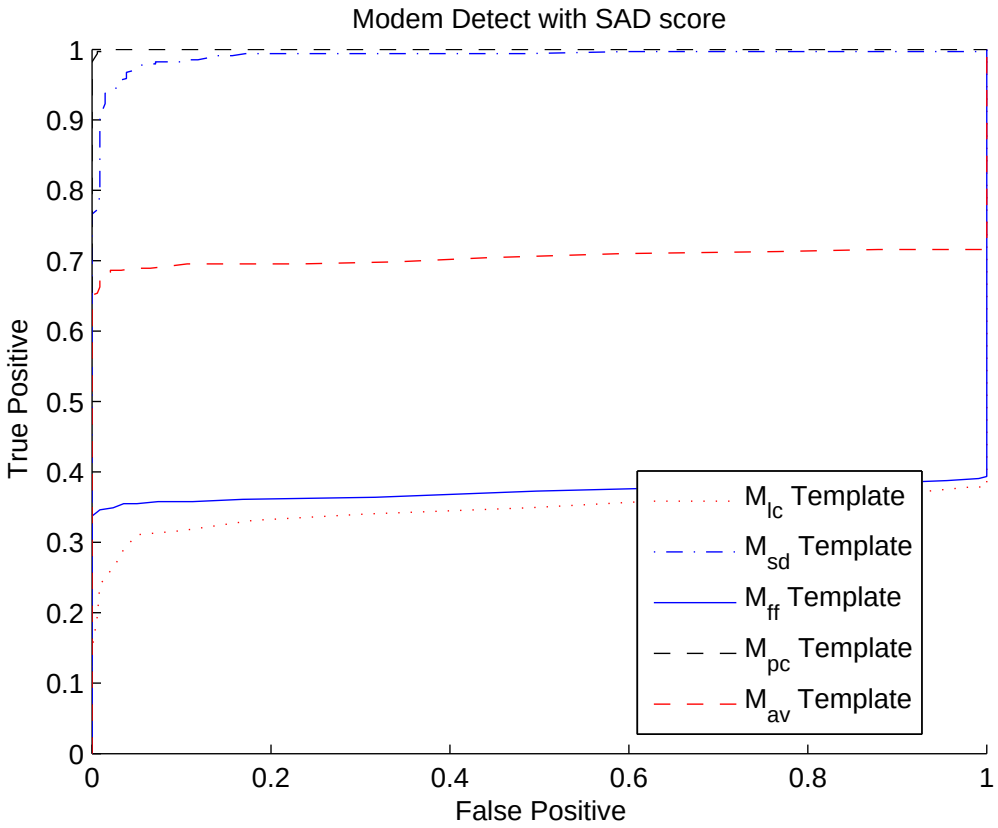

Figure 25. ROC graph of modem templates with SAD

Table VI. Ground template fit weighted score summary for Fisherman Flats.

Template	excellent	good	poor	unacceptable
G_{ff}	6	2	2	0
G_{lc}	0	2	4	4
G_{sd}	4	5	1	0
G_{pc}	0	9	1	0
G_{av}	0	4	4	2

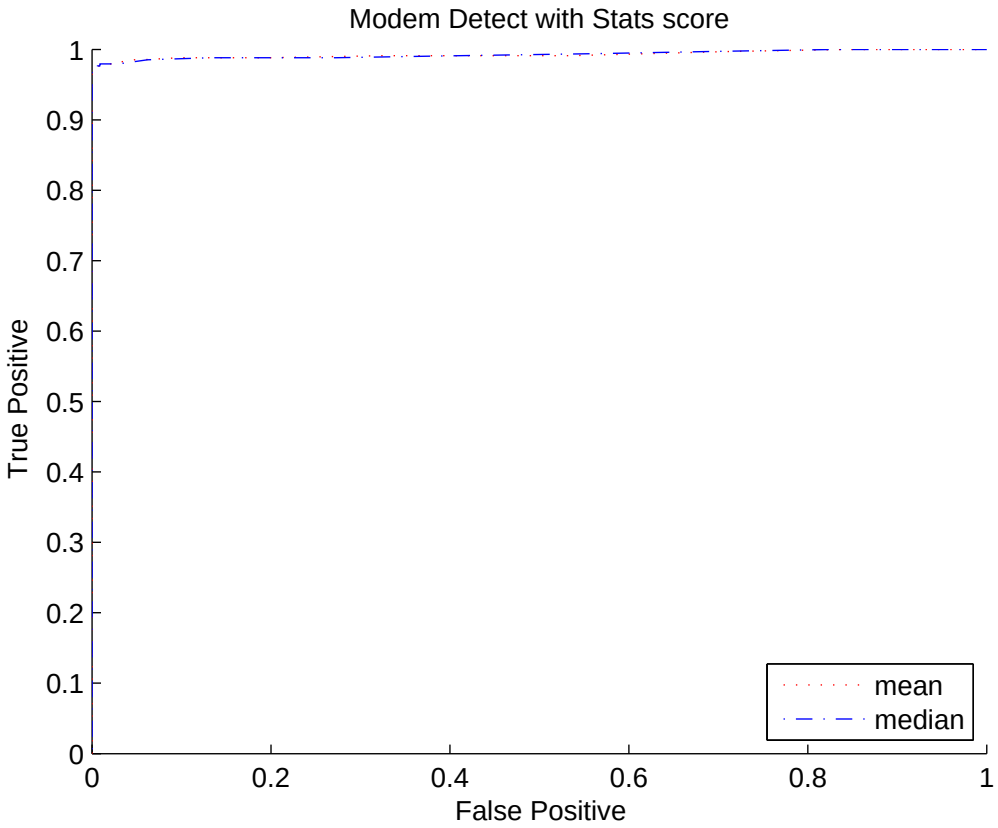

Figure 26. ROC graph of modem detection with statistic scores

Table VII. Ground template fit weighted score summary for Lovers Cove.

Template	excellent	good	poor	unacceptable
G_{ff}	0	1	2	7
G_{lc}	0	0	2	8
G_{sd}	0	0	2	8
G_{pc}	0	0	1	9
G_{av}	0	1	2	7

Table VIII. Ground template fit weighted score summary for San Diego.

Template	excellent	good	poor	unacceptable
G_{ff}	3	7	0	0
G_{lc}	0	4	3	3
G_{sd}	5	5	0	0
G_{pc}	1	9	0	0
G_{av}	0	0	10	0

Table IX. Ground template fit weighted score summary for Panama City.

Template	excellent	good	poor	unacceptable
G_{ff}	3	6	1	0
G_{lc}	0	3	6	1
G_{sd}	2	7	1	0
G_{pc}	0	4	6	0
G_{av}	0	2	7	1

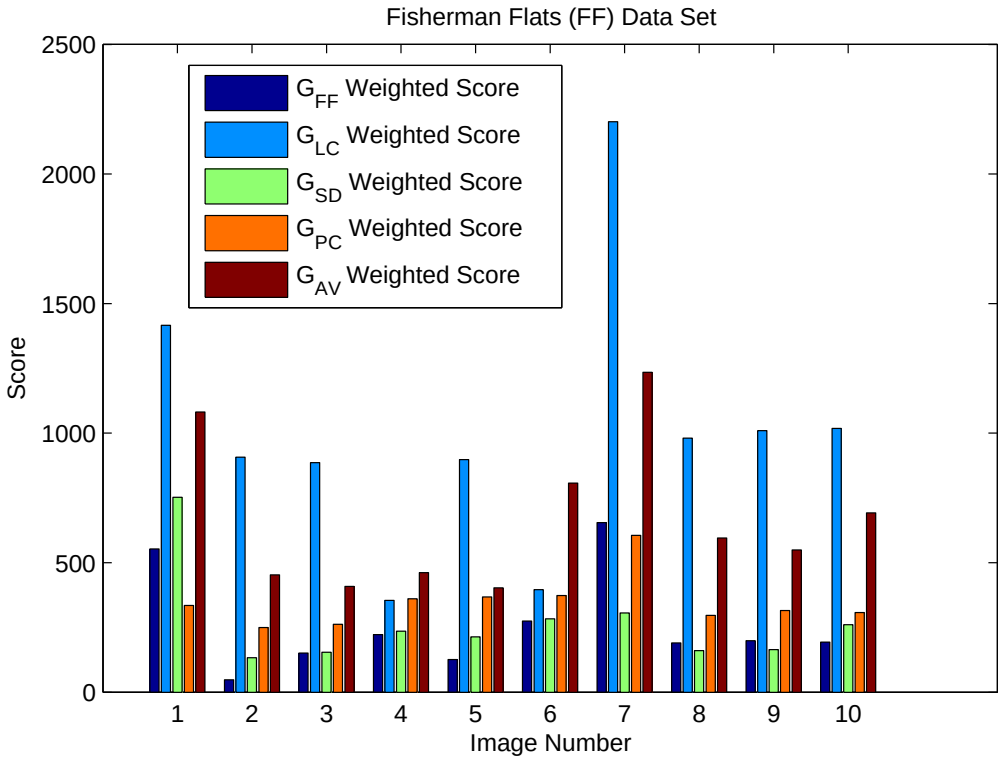

Figure 27. Plot of weighted score of five template fit to Fisherman Flats dataset

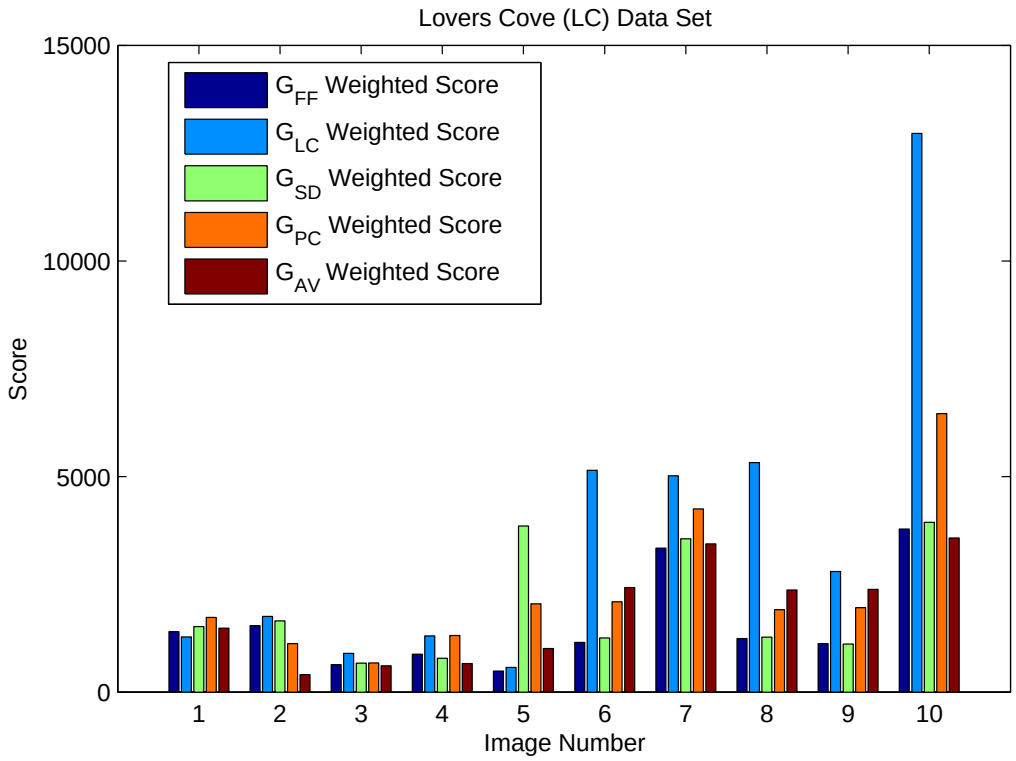

Figure 28. Plot of weighted score of five template fit to Lovers Cove dataset

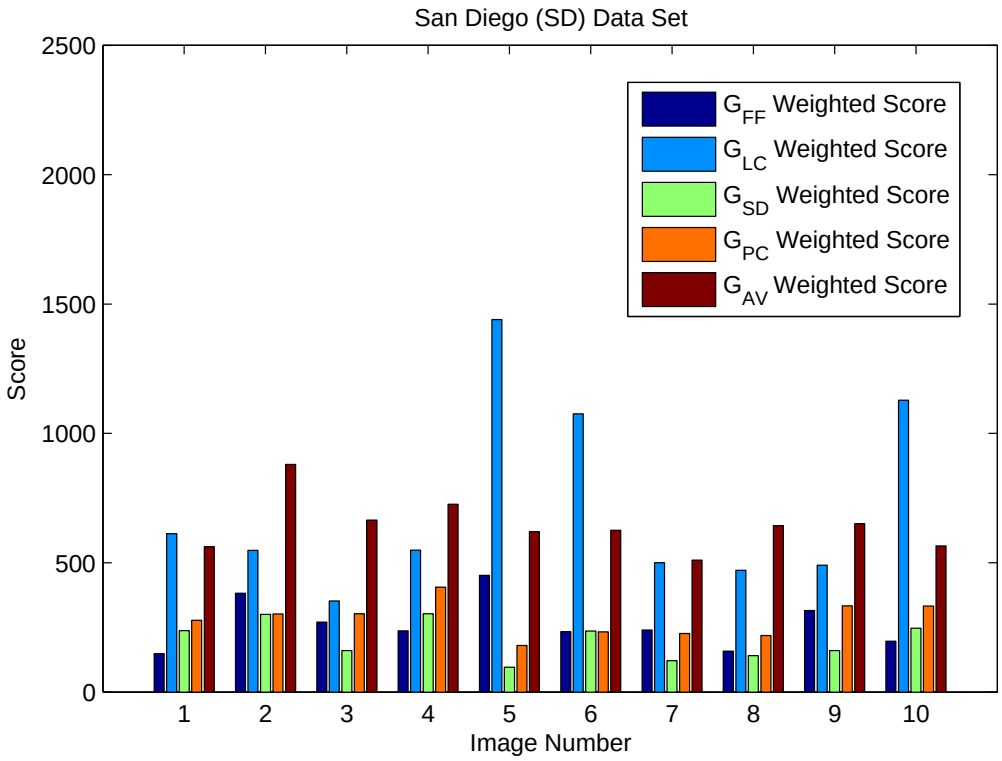

Figure 29. Plot of weighted score of five template fit to San Diego dataset

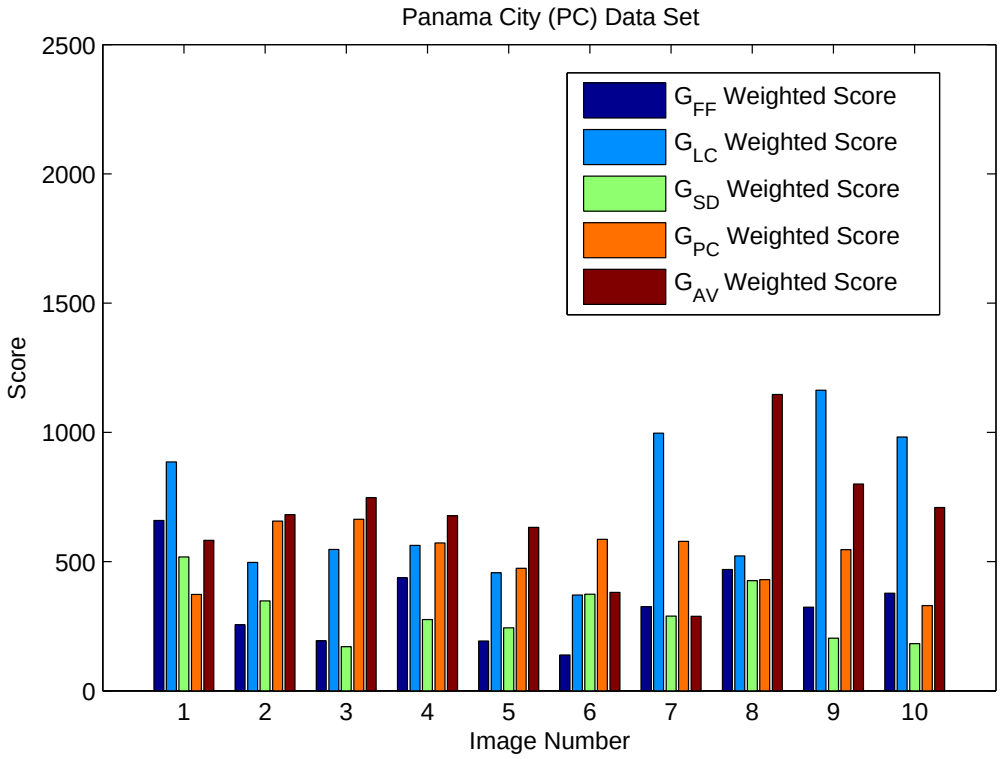

Figure 30. Plot of weighted score of five template fit to Panama City dataset

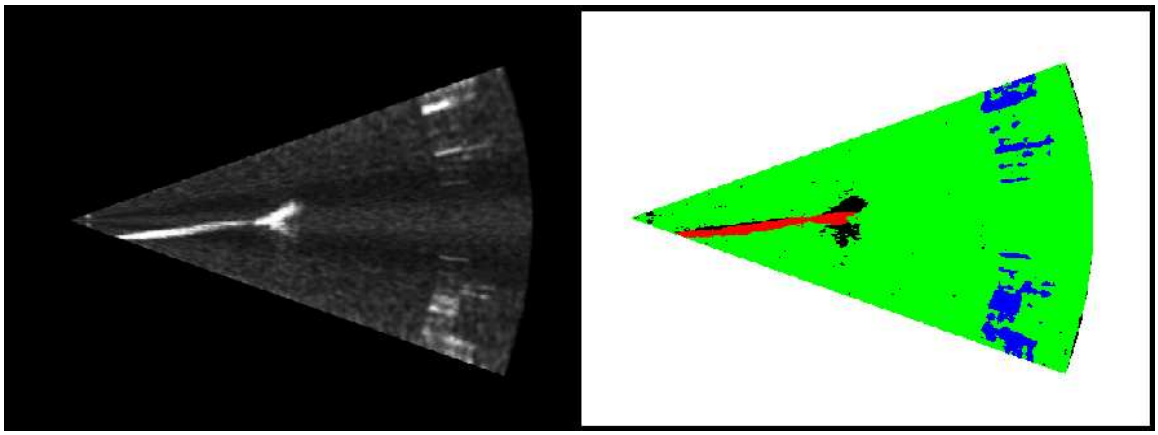

Figure 31. Left: Sonar Image with obstacle protruding from ground. Right: Processed Sonar Image.
Green=background, Red=ground, Blue=pulse noise, Black=unknown object(approaching rock wall)

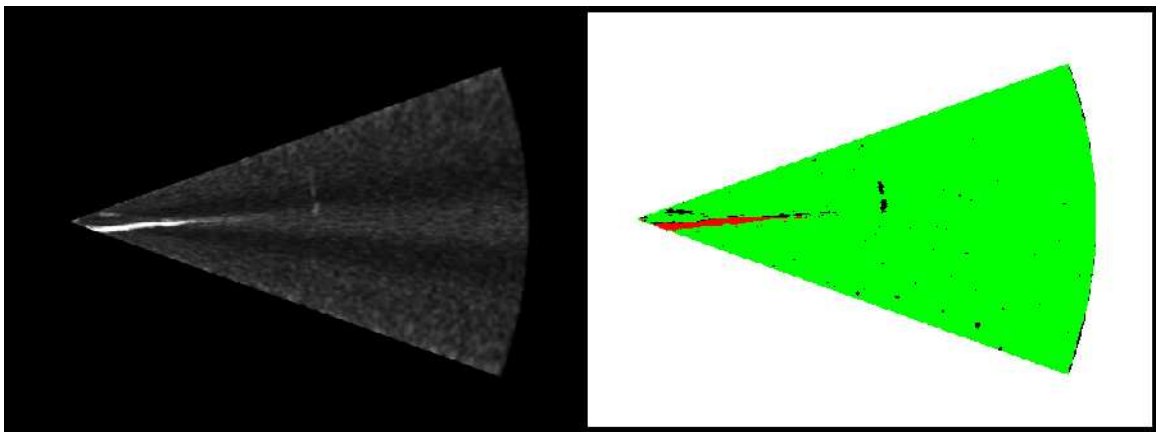

Figure 32. Left: Sonar Imagewith faint obstacle approaching. Right: Processed Sonar.
Green=background, Red=ground, Black=unknown object

V. DISCUSSION

A. ACOUSTIC IMAGE MODEL

The results show that the models explain the observed patterns fairly well. The background models fit to over 96% of the pixels within 3σ, which were exploited to define a good probability function for segmentation into background and other artifacts. There were some issues with the background classification experiments. The results of the background ROC curves inaccurately portrayed the performance of the background segmentation, caused by errors in annotation. This was a result of techniques used to quickly create annotations. This issue only affected the background ROC test because annotated features contained a noticeable amount of background noise.

The template-based pulse noise detection performs well in the top portion of the image using the Pt_{SD} template matched with SAD. This was able to classify images containing pulse noise correctly 92% of the time with less than a 3% false positive rate. The same template scored with NCC performed substantially worse with a detection rate of 60% with 50% false positives. Overall for detection of top-pulse noise the templates scored with SAD performed much better than those scored with NCC and the three templates Pt_{SD}, Pt_{av} and Pt_{pc} that performed well scored with SAD performed the worst when scored with NCC. None of the bottom-pulse template-based detections performed better than the top-pulse. We suspect this was because the ground was interfering with the template matching, increasing the false positives. The Pb_{pc} template scored with SAD performed the best out of all the bottom-pulse template experiments. This template had an 80% detection rate with less than 7% false positives. While most of the templates did not perform well, a few were able to detect images with pulse noise. Template matching is a reliable way to detect pulse noise, but the right template is important. Using the mean and median of small areas was not reliable in detecting any of the pulse noise. For bottom-pulse

noise both median and mean were 50% detection to 50% false positives. Up-pulse noise was slightly better. Future work should investigate using a Bayesian classifier to get a single score from the up-pulse and bottom-pulse test.

It should be noted that the site-invariant template (Pt_{av}) did not perform better than third in any of the pulse noise detection experiments. We suspect when the Pt_{av} template was created by averaging the templates together, the features that make a good template were muted.

A 92% detection rate for pulse noise is acceptable. Pulse noise is not as common as other noise and it effects a relatively small area. A image misclassified as not having pulse noise will not have a substantial effect on latter processes. The pulse area will appear as three to five extra objects in an image.

Modem noise detection proved robust with any of the three techniques. With Template-based modem noise detection with NCC, all five of the modem templates (M_{ff}, M_{lc}, M_{sd}, M_{pc}, M_{av}) accurately detected images with modem noise more than 80% of the time with less than 5% false positives. The M_{pc} template with the SAD scoring performed the best with 100% detection and 1% false positives. Calculating the mean and the median were both very reliable at detecting images with modem noise 98% of the time with less than 2% false positive.

Being able to accurately identify images with modem noise is important because the noise effects such a large area of the image. If an image is misclassified as not having modem noise then the large unknown area will be classified as objects and possible collisions. During the ground-template-fitting stage a ground model will be fit to each one of these objects. This will increase the run time of the ground-template-fitting stage exponentially.

The method for ground position and orientation estimation with a generic template succeeded in the majority of the cases. The G_{ff} and G_{sd} templates performed the best at multiple sites. One interesting result was that the G_{ff}, G_{sd} performed the best at their corresponding site, but G_{pc}, G_{lc} did not. Also, the G_{av} was one of the

poorest performing templates over all the sites. Using template matching to find the ground worked well at specific sites but did not do well at describing the ground at the Lovers Cove site. The ground at this site is very rocky and dynamic and difficult to describe with a template.

B. SOFTWARE DEVELOPMENT

Developing the software architecture in two stages was very successful. Most development time was spent testing different classification techniques. Once these were complete the UML digram was created in an afternoon. In just over a week 80% of the C++ code was written and debugged. Currently implemented is the background classification, pulse noise detection/classification and modem noise detection/classification. These sections have been debugged and tested and the results match those of MATLAB. Other functions needed such as polar to Cartesin are also implemented. Taking advantage of object-orientated concepts such as inheritance testing in the horizontal plane using these techniques has already begun. Tasks that are still remaining are implementing the ground template matching which is 50% complete and creating benchmark functions that use real-time clock to get sub-millisecond timing. The development schedule would not have been possible using a single programming language.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSIONS

This thesis tested the quality of models for detecting and locating acoustic image artifacts across a range of ocean environments. The four artifacts modeled were all well detected and described. Experiments demonstrated that images containing pulse noise and modem noise can be reliably classified. This is critical because in some situations devices like acoustic modems are required. For example, in multi-vehicle operations acoustic modems are a common way for vehicles to communicate over ranges greater than a few meters. In scenarios with a constant bottom type, using template matching was a succesful way to describe the ground. With the ground described vehicle state information can now be estimated.

Reliable, robust detection is critical for navigation and obstacle avoidance based on FLS data analysis. In comparison to prior methods, this can be performed in still images and entirely without knowledge of metadata, as opposed to relying on temporal information, smoothing and consistency.

There was one experiment that was inconclusive. The test data set to quantify the background classification were annotated incorrectly. A new set of images need to be annotated and this test rerun. The methods described in this thesis are scheduled to be tested live on-board a vehicle early January 2009. To meet the testing deadline the conversion to C++ needs to be finished and these methods applied to the horizontal plane. Once testing is complete future work will include decreasing process time and refining the current models. For example some of the templates performed much better than others, but it is not known why. It is unclear what makes a good template for the different artifacts. Future work will also include creating models for artifacts that are not currently described. In conclusion this thesis describes robust methods to detect and classify common artifacts in FLS images.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1] T. Allen, A. Buss, and S. Sanchez. Assessing Obstacle Location Accuracy in the REMUS Unmanned Underwater Vehicle. *Proceedings of the 2004 Winter Simulation Conference*, pages 940–948, 2004.

[2] M. Benjamin. Interval Programming: A Multi-Objective Optimisation Model for Autonomous Vehicle Control. *PhD. Dissertation, Brown University, Providence, RI*, May 2002.

[3] M. Dolbec. Velocity Estimation Using Forward Looking Sonar. *M.S. Thesis Naval Postgraduate School*, 2007.

[4] J. Doucette. Navigation Autonomous Underwater Vehicles. http://www.whoi.edu/oceanus/viewImage.do?id=19010&aid=9207. [Accessed December 10,2008], 2006.

[5] E. Dura, J. Bell, and D. Lane. Superellipse tting for the classication of mine-like shapes in side-scan sonar images. *Proc. MTS/IEEE Oceans, Conf. and Exhibition*, pages 23–28, 2002.

[6] J. Folkesson, J. Leonard, J. Leederkerken, and R. Williams. Feature Tracking For Underwater Navigation using Sonar. *IEEE/RSJ International Conference on Intelligent Robots and Systems Proceedings*, 2007.

[7] W. L. J. Fox, , J. Hsieh, and C. Polwarth. Segmentation of images from an acoustic lens sonar. *OCEANS '04. MTTS/IEEE TECHNO-OCEAN '04*, 4:2029–2035, 2004.

[8] J. Han and A. Asada. Acoustic Counting Method of Upstream Juvenile Ayu Plecoglossus altivelis by Using DIDSON. *Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, pages 459–462, 2007.

[9] R. M. Haralick and L. Shapiro. *Computer and Robot Vision* , volume 2. Addison-Wesley, 1993.

[10] D. Horner and O. Yakimenko. Recent Developments for an obstacle Avoidance System for a Small AUV. *paper presented at IFAC Conference on Control Applications in Marine Systems, Bol island Brac, Croatia*, 2007.

[11] Hydroid. REMUS 100 Specifications. http://www.hydroidinc.com/100spec.html. [Accessed October 25,2008] , 2008.

[12] F. J., J. Leederkerken, R. Williams, A. Patrikalakis, and J. Leonard. A Feature Based Navigation System for an Autonomous Underwater Robot. *Springer Tracts in Advanced Robotics*, 42/2008,:105–114, 2008.

[13] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy. Statistical model and genetic optimization: application to pattern detection in sonar images. *IEEE International Conference Acoustics, Speech and Signal Processing Proceedings*, 5:2741–2744, 1998.

[14] J. Nelder and R. Mead. A Simplex Method for Function Minimization. *J. Computer*, 7:308–313, 1965.

[15] Y. Petillot, I. Tena Ruiz, and D. M. Lane. Underwater vehicle obstacle avoidance and path planning using a multi-beam forward looking sonar. *IEEE Journal of Oceanic Engineering*, 26:240–251, 2001.

[16] I. Quidu, P. Malkasse, G. Burel, and P. Vilbe. Mine classication using a hybrid set of descriptors. *Proc. Oceans MTS/IEEE Conf. and Exhibition*, 1:291–297, 2000.

[17] S. Reed, Y. Petillot, and J. Bell. An Automatic Approach to the Detection and Extraction of Mine Features in Sidescan Sonar. *IEEE Journal of Oceanic Engineering*, 28:90–105, 2003.

[18] S. Reed, Y. Petillot, and J. Bell. Automated approach to classication of mine-like objects in sidescan sonar using highlight and shadow information. *Radar, Sonar and Navigation Proceedings*, 151:48–56, 2004.

[19] S. Reed, Y. Petillot, and J. Bell. Model-Based Approach to the Detection and Classification of Mines in Sidescan Sonar. *Applied Optics*, 43:237–246, 2004.

[20] B. Russell, A. Torralba, K. Murphy, and W. T. Freeman. LabelMe: a database and web- based tool for image annotation. *IJCV*, 2008,.

[21] H. Sekkati and S. S.Negahdaripour. 3-D Motion Estimation for Positioning from 2-D Acoustic Video Imagery. *Lecture Notes in Computer Science*, 4478:80–88, 2007.

[22] S.-C. Yu, T.-W. Kim, A. Asada, S. Weatherwax, B. Collins, and J. Yuh. Development of real-time acoustic image recognition system using by autonomous marine vehicle. *Ocean Engineering*, 35:90–105, 2008.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center

2. Dudley Knox Library
   Naval Postgraduate School