

Ground Truth Tool for Verification of Small Boat Tracking

by

Zachary Rovig

A Thesis Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2021

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Master's Committee, we certify that we have read the thesis prepared by Zachary Rovig, titled Ground Truth Tool for Verification of Small Boat Tracking and recommend that it be accepted as fulfilling the thesis requirement for the Master's Degree.

Dr. Lloyd LaComb

Date: _____

Dr. Pierre Blanche

Date: _____

Dr. Amit Ashok

Date: _____

Final approval and acceptance of this thesis is contingent upon the candidate's submission of the final copies of the thesis to the Graduate College.

We hereby certify that we have read this thesis prepared under our direction and recommend that it be accepted as fulfilling the Master's requirement.

Dr. Lloyd LaComb
Master's Thesis Committee Co-Chair
College of Optical Sciences

Date: _____

Dr. Pierre Blanche
Master's Thesis Committee Co-Chair
College of Optical Sciences

Date: _____

ACKNOWLEDGEMENTS

I want to thank my advisor, Dr. Lloyd LaComb for making this whole process enjoyable and being available to me whenever I needed help. I also want to thank Dr. Pierre Blanche for accepting me into his lab as an undergraduate and helping me build essential skills in optics. Finally, I want to acknowledge my classmates and say thank you for all those hours we spent together in the optics lounge ironically laughing about our small failures, which seemed so important at the time.

DEDICATION

To my parents, Chris and Christy . Thank you for all your love and support.

Table of Contents

Abstract	7
1 Introduction	8
2 Background and Related Work	9
2.1 Multispectral and Polarimetric Approaches	9
2.2 Automatic Tracking Algorithms.....	14
3 Multispectral/Hyperspectral Polarimetric Imaging	17
3.1 Background	17
3.2 Statements on Optical Design	20
4 Automatic Tracking Recognition (ATR)	25
4.1 ATR Background	25
4.2 Data Fusion.....	25
4.3 Kalman Filtering.....	26
4.4 Reed – Xiaoli Anomaly Detection	29
4.5 Detection Results	30
5 Verification Via Ground Truthing	32
5.1 Ground Truth Data Background	32
5.2 Developed Ground Truth Tool	34
6 Conclusions	37
APPENDIX A – TABLE OF ACRONYMS	39
References	40

List of Figures

- 1 Image of a tank in thermal equilibrium captured with infrared camera. (Left) intensity image. (Right) Degree of linear polarization.(Courtesy: David Cenault, Polaris).¹
- 2 Reflection coefficient versus angle of incidence (from surface normal) for an air-water interface
- 3 Block diagram of focusing lens and detector
- 4 Plot of Atmospheric Opacity (%) versus Wavelength (m) with system spectral regions highlighted
- 5 Plan view (left) and cross sectional view (right) on Sony's IMX250MZR sensor. Each pixel's polarizing filter (C) is coated with an anti-reflective layer (B) and is positioned between the microlens (A) and the light sensitive photodiode (E). Image from Sony Corporation²
- 6 Preliminary Zemax optical model of the five band imaging system
- 7 Preliminary Data Fusion and Analysis Model (courtesy of Dr. Mark Neifeld)
- 8 Tracking algorithm system architecture (credit: Dr. Abhijit Mahalanobis)
- 9 ROC curves from RXD results (credit: Dr. Abhijit Mahalanobis)
- 10 Single Frames from each polarization channel (1-4) and same spectral band
- 11 Screenshot of the GUI from the developed Ground Truth App
- 12 Screenshot of the GUI from the developed Ground Truth App with Tab selection highlighted
- 13 Screenshot of the "Review" tab after point selection is finished
- 14 Screenshot of the tables on the "Point Selection" tab (left) and "Review" tab (right)

Abstract

The U.S. Navy is in need of an imaging system to perform new defense strategies such as the Find-Fix- Track-Target-Engage-Assess (F2T2EA) to minimize threats from fast attack crafts in geopolitical hotspots. F2T2EA requires near-instantaneous detection and tracking of fast attack crafts in littoral waters. In this paper, we present a multispectral polarimetric imaging (MSPI) system, automatic target detection and tracking algorithm, and custom ground truth data tool for tracking verification. A review of studies using multispectral imaging for object detection and automatic tracking algorithms was done and is shared in the early sections of this paper. The designed MSPI system uses dichroic mirrors along with division of focal plane (for visible and NIR bands) and division of amplitude (for SWIR, MWIR, and LWIR bands) to separate the spectral bands and polarization components, respectively. The tracking algorithm adopts techniques from Kalman filtering and Reed – Xiaoli anomaly detection. Results from the developed tracking algorithm are reported. The results show that including single frame spatial analysis as a complement to motion detection across several frames improves the detection and tracking performance of the algorithm. The custom ground truth app allows manual tracking of up to twelve objects, frame incrementation, and linear interpolation of between selected points.

1 Introduction

Small fast-moving boats, such as the Fast Attack Craft (FAC) and the Fast Inshore Attack Craft (FIAC), are constantly threatening the assets of the US Navy. Various US adversaries have recently engaged in aggressive interactions using FACs. These encounters occur in geopolitical hotspots such as the Straits of Hormuz³ and off the coast of Somalia.⁴ The incidents occurring in the Straits of Hormuz involved multiple Iranian FIACs approached a US guided-missile destroyer at high speeds with their crew-served weapons manned. The crew of the US guided-missile destroyer tried multiple times to warn off the FIACs with radio communications, sirens, and the ship's whistle to no avail. The US guided-missile destroyer was forced to resort to firing warning shots from the ship's 0.50 caliber guns when the FIACs came within 900 yards, finally causing them to break off. This is one of several events that have occurred in the Straits of Hormuz in the last few years.

Further, Iran and North Korea have some of the largest numbers of FAC's in operation. North Korea alone operates more than 300 FAC's,⁵ while Iran has developed a fleet of 'swarm boats' to harass vessels in the heavily congested littoral waters of the Persian Gulf. The US Navy is introducing new strategies such as Find-Fix-Track-Target-Engage-Assess (F2T2EA) and applying advanced solutions to reduce the threat of FACs in littoral waters. Using multispectral electro-optical imaging to quickly spot, locate, and monitor small fast attack crafts is a key force multiplier for the US Navy. Current electro-optical systems are usually configured for ground-based operations and do not consider the effects of ocean waves and wakes, which may be cause for confusion in identification and tracking systems. The Navy needs an improved electro-optical/infrared (EO/IR) imaging system for detection, and identification of small, fast, agile boats in the littoral theatre. Better protection for US naval forces can be provided with a multispectral EO system with im-

proved discrimination in ocean waters. Improved discrimination and identification can be obtained with the ability to use the wakes generated by fast attack boats.

The overarching goal of this project is to provide an imaging and tracking system to improve discrimination of FAC/FIACs in littoral waters. To accomplish this goal, a prototype multispectral polarimetric imaging system and corresponding target recognition algorithms were developed and are presented, generally, in this paper. These designs are presented on behalf of my team at University of Arizona and TIPD: Dr. Lloyd LaComb, Dr. Pierre Blanche, Dr. Amit Ashok, Dr. Mark Neifeld, and Dr. Abhijit Mahalanobis. My contribution to the project lies within the development of a tool to establish ground truth data, which can be seen in greater detail in section 5.2.

2 Background and Related Work

2.1 Multispectral and Polarimetric Approaches

Various studies have focused on small boat detection in the visible and IR spectral bands. Of these studies, some notable DoD-authored papers provide a basis for design considerations and methodologies. The methods used in these studies include imaging system design using MWIR and LWIR bands in night-time detection⁶ and required sensor parameters necessary for boat detection when using visible and MWIR spectral bands.⁷ The Naval Postgraduate School has also produced some research on using polarization to improve contrast for boat detection in ocean settings. One study highlights the potential improvements and complications of including polarization measurements.⁸ Further effects of polarization on land-based image contrast and measurement techniques are described from Snik and Craven-Jones et al¹

The paper focused on nighttime boat detection, by Buss and Ax,⁶ compared the performance of MWIR and LWIR spectral bands. Although the requirements and specifications outlined are

different from this project, the paper still presents a sound methodology. The spectral bands candidates were determined viable based on the long detection range (several miles). Buss and Ax found that thermal infrared bands were going to be the only viable options since the reflected-light spectra (visible to short-wave IR) show poor target contrast on the ocean surface and nighttime/low visibility conditions cannot be overcome and active illumination was not an option. Thermal imaging provides potential for a higher target contrast over many environmental conditions. The hull of ships and many materials commonly found on marine vessels have a higher emissivity than the water and reflected sky. Furthermore, any heat sources aboard the vessel such as people, power sources, engine signatures, etc. will provide additional signal contrast than the hull of the vessel itself.

For nighttime boat detection, Buss and Ax considered five spectral bands in the thermal IR: broadband MWIR (3.4 – 5.0 μm), narrow MWIR (3.4 – 4.2 μm), notched MWIR (3.4 – 4.15 μm and 4.6 – 4.8 μm), cooled LWIR (8.0 – 10.5 μm), and uncooled LWIR (8.0 – 12.0 μm). It is important to note that the narrow MWIR cuts off at the beginning of the CO_2 absorption line and the notched MWIR uses a filter to cut out the CO_2 absorption from 4.2 – 4.6 μm . The cooled vs uncooled LWIR refers to the detector type that was used with each band. The uncooled detectors have lower sensitivity than the cooled detectors, but they boast low prices and low overhead achieved by eliminating the cooler. Four spectral bands, out of the five originally considered (broadband MWIR was excluded), were simulated through the Navy Littoral database at a target range of 12 miles and atmospheric transmissions were observed. It was shown that the LWIR show higher transmission in their highest transmission cases, but faster drop-offs in transmission compared to the narrow and notched MWIR bands. These simulation results were considered in the design of the imaging system developed for our project.

Krapels and Driggers⁷ used a target discrimination criterion to determine the requirements for successful sensor realization. The criterion was used to characterize the performance of existing infrared sensors and in the design of new conceptual sensors. Daytime and nighttime performances were explored in visible and MWIR spectral bands, respectively. The experimental approach involved developing a representative target set for small watercrafts and collecting signatures for creating a target identification perception experiment. Trained observers then participated in the identification experiments and the results of these experiments were used to determine the sensor discrimination criteria. The boat signatures were collected during the day in the MWIR (3 – 5 μm), LWIR (8 – 12 μm), and visible bands. However, only the MWIR and visible bands were processed for this article. A total of 576 signatures were collected and processed consisting of 12 watercrafts viewed at 12 aspect angles in each of the two wavelengths bands.

The results from the observer perception experiment were presented as an average probability of identification across all the observers. The probabilities were determined by comparing the cycles on target value to the N_{50} and V_{50} discrimination criterion. The 50% probability of small boat identification (N_{50}) was 4.0 cycles on target in the visible band and 2.8 cycles on target in the daytime MWIR band. For primary metric used in the Army and Marine Corps, the 50% probability of boat identification (V_{50}) is 14.0 cycles on target in the visible and 10.6 cycles on target in the daytime MWIR. In this case, “cycles on target” was converted from cycles per milliradian by multiplying the limiting frequency by the target angular subtense.⁷ The discrimination criteria presented in this paper is important to consider when choosing a sensor for the marine target tracking system.

A study done by Cooper et al.⁸ at the Navy Postgraduate School (NPS), measured the effects of polarization on target-to-background contrast measurements. The data was collected using MWIR

(3 – 5 μm) and LWIR (8 – 12 μm) with polarizers mounted in either the vertical or horizontal positions. The polarization data was taken in pairs with the polarization axis being rotated manually between image pair collection. The collected data showed that when the polarizer is oriented in its horizontal position, the apparent sea radiance is more effectively reduced. Reducing the sea radiance around the target increases the overall contrast between the target and background. This result led Cooper et al. to infer that the sea surface radiance is primarily vertically polarized, while the target ships were much less so.

This study also researched the impact of different solar angles by capturing images at different times throughout the day from 9am – 9pm. The goal was to show the contribution of solar reflections to the overall degree of polarization of the sea surface. It was determined that the morning periods provided a more obvious improvement in target-to-background contrast. The horizontal polarizer also enhances the interface contrast between the sky and the sea, helping to locate the horizon line within the field of view. The results of the study show that man-made targets did not show a significant polarization preference, the sea background was predominantly vertically polarized, and horizontal polarizers provide a higher contrast between the targets and the sea-background. It was also found that LWIR provides a higher horizon and target-to-background contrast than the MWIR band. It is expected that the emission polarization will be more prominent in the LWIR band and the reflection polarization more so in the MWIR band.

Polarization data provides information that cannot be obtained using spectral data alone. Polarization properties depend on the surface roughness, shape, and orientation of the object. For land-based detection, object-to-background contrast can be significantly improved for objects in thermal equilibrium with their environments.¹ Figure 1 shows an example of thermal IR cameras and two images, one captured in intensity and the other captured with a degree of linear polariza

tion.

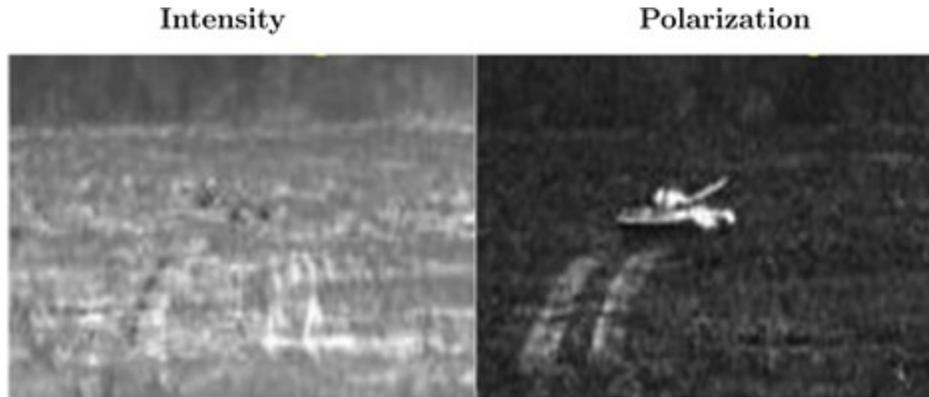


Fig 1 Image of a tank in thermal equilibrium captured with infrared camera. (Left) intensity image. (Right) Degree of linear polarization.(Courtesy: David Cenault, Polaris).¹

Polarization measurement techniques are shared by Snik and Craven-Jones et al based on the results from a workshop on “Polarimetric Techniques & Technology”.¹ Measurement techniques that are designated as being in the spatial domain are considered the most straightforward. These techniques include separation of polarization using a polarizing beam splitter (PBS) prior to the sensors and using microgrid polarizers to separate each polarization state at each pixel on the detector. PBS systems can be designed to yield four beams to provide measurements of the Stokes parameters (shown in equation 1). These systems allow for instantaneous polarization measurements but can be bulky and difficult to boresight each sensor. Division of focal plane methods through microgrid polarizers create grids of pixels, each pixel (or block of pixels) measuring a different degree of linear polarization. Focal plane division also provides instantaneous polarization methods and removes the issue of bore sighting multiple sensors, however, at the cost of reduced spatial resolution. Both techniques discussed here are considered in the prototype design shown in section 3.2.

2.2 Automatic Tracking Algorithms

Tracking algorithms can take on many different forms. This project focused mainly on Kalman filtering-based approaches as they are often the best selection for video-based tracking. Kalman based tracking methods including successive clustering,⁹ optical flow,¹⁰ and mean shift tracking¹¹ are discussed in the following paragraphs. A more in-depth explanation of Kalman filters can be found in section 4.3. More complex automatic tracking algorithms can be created using machine learning and neural networks.¹² These automatic trackers typically require extensive training data and excessive computational power. There is no readily available training data for this project so automatic tracking algorithms using machine learning were not chosen but are included here for completeness.

In a study done by Bloisi,⁹ a visual detection method using Haar-like features was applied to detect and track boats on the water surface. Harr-like features are digital image features, adapted from Harr basis functions, that are used in object recognition. These features are adjacent, same-size rectangular regions that categorize subsections of an image. A Harr-like feature based system provides ad hoc domain knowledge more effectively than finite training data and operates much faster than a pixel-based system.¹³ Using a Harr-like feature based approach requires little computational power and therefore provides real-time results. The Haar-like classifier was trained using 4000 images not containing boats and 1500 images showing different types of boats. The detection algorithm highlights the horizon and set a bounding box around each boat. Bloisi et al. found that limiting the detection area to below the horizon line decreases the number of false positives in the system. Finally, an additional weak classifier was established to recognize boat wakes and further increase the detection rate (DR) and reduce the false alarm rate (FAR). The final DR and FAR for

this study are 0.928 and 0.251, respectively, surveyed across a set of 100 randomly chosen images.

In another study done by Bloisi,¹⁰ a video surveillance system (ARGOS) for boat traffic monitoring in Venice, Italy was presented. ARGOS is a four-camera system that surveys a stretch of canals 24 hours a day, 7 days a week. The tracking module used with ARGOS relies on optical flow, blob formation, and clustering. Optical flow considers the correlation between two consecutive frames to handle under and over detection, that is detecting fewer and more objects, respectively, than the actual number of objects. Optical flow finds the features that are present in both frames and creates a motion vector for each of the objects. An optical flow map is then created consisting of colored points with colors indicating moving direction. Clustering these optical flow maps helps improve detection effectiveness. Under detection is avoided when two boats are close but moving in opposite direction and over detection is avoided by discarding blobs with a small number of optical flow points. One of the limitations of the optical flow method occurs when a boat turns around itself. This could result in the motion vector of the boat drastically changing between frames and the boat being discarded for some frames. The tracking method presented in this study resulted in a high-count accuracy (0.941), an average position error of about 5 m, and average velocity error of about 1 km/hr.

Comaniciu and Ramesh et al¹¹ developed a tracking algorithm using isotropic kernels. The target is first represented by an ellipsoidal region within the image. An isotropic kernel then assigns smaller weights to pixels farther from the center of the ellipse. The target localization procedure then starts from the position of the target in the previous frame and searches in the surrounding pixels for the center of the target. Assumptions are made that require that the target position does not change drastically from the initial frame, which is most often valid for consecutive frames. This kernel-based target localization method was then integrated with a Kalman filtering framework.

The mean shift procedure finds the root of the gradient as a function of location. This corresponds to a similarity surface between target locations. When the Kalman filtering framework is applied the similarity surface is normalized and represented as a probability density function. Results were shown involving the tracking of a person's hand. The algorithm was able to track the target hand even in the presence of complete occlusion by a similar object (the person's other hand). This result is very promising for our project as there can be many instances where boats we want to track cross paths and/or become occluded by other boats.

Over the last few years convolutional neural networks (CNN) have emerged as a powerful machine learning tool. CNNs have made significant improvements in computer vision problems, including object detection. An ATR system, developed by Nasrabadi,¹² based on deep convolutional neural networks (DCNN) was used to detect targets in forward looking infrared (FLIR) scenes. Two scenes used for system evaluation were ground-based and taken during the summer in the Arizona desert and in the spring in central California. Targets in these datasets were shot at distances ranging from 688 to 3403 meters. The DCNN was made up of two networks: the first network is responsible for target detection and feed the results to the second network where targets are classified into their target types and false alarms from the first network are rejected.

The DCNN was trained using a dataset of approximately 5,000 images for 100 epochs (passes on entire training dataset) in batches of 200. Training was an iterative process based on the performance of the network on the validation dataset. The DCNN was sent through three training processes before being evaluated on the test scenes in Arizona and California. Detection results were obtained before and after the second network rejected the false alarms received from the first network. Prior to false alarm rejection the DCNN achieved a probability of detection of 88% for 0.35 false alarms per frame in the California dataset and 78% for 0.5 false alarms per frame in

the Arizona dataset. After false alarm detection in the second network the detection probabilities significantly improved, reaching 99.8% for 0.35 false alarms per frame in the California dataset and 97.9% for 0.5 false alarm per frame in the Arizona dataset. It was expected that the Arizona dataset would prove to be the more difficult test for the DCNN as the images were taken in FLIR and the target-to-background temperature difference was less than that in the California datasets. This ATR system using DCNNs was shown to improve target detection and recognition probabilities when compared to previous state-of-the art systems. However, since this system was only evaluated on pre-collected datasets, there is no way to know how the tracking system would perform with real-time data collection and analysis.

3 Multispectral/Hyperspectral Polarimetric Imaging

3.1 Background

Multispectral polarimetric imaging provides images with multiple spectral bands and polarization information. Multispectral imaging (MSI) and Hyperspectral imaging (HSI) are described and a comparison between the two techniques is made in the following paragraphs. Further, the addition of polarization components and how that information can complement the spectral data will be discussed. MSI describes methods of spectral imaging where the obtained images contain multiple (as many as tens) spectral bands. Often the spectral bands of interest are not fully contained within the visible spectrum, covering parts of the ultraviolet (UV) and infrared (IR) regions. As shown in section 3.2, the multispectral imaging system designed for this project contains five spectral bands ranging from the visible to thermal IR bands. The limited number of spectral bands results in a low spectral resolution and may not collect all the information needed for small object detection. In this case, small objects are relative to the ground sampling distance (GSD) of the imaging system.

A large GSD can result in the objects of interest being embedded in a single pixel and, therefore, difficult to resolve with limited MSI data. The superior spectral resolution of HSI can improve detection ability for both sub-pixel and larger objects.

HSI sensors make use of many contiguous spectral bands to expand on the capabilities of multispectral imaging sensors. The high spectral resolution allows HSI sensors to discriminate between more subtle objects and features, when compared to multispectral imaging sensors. It is intuitive to think of HSI as an extension of multispectral imaging, but in reality, it is not that simple. Chang¹⁴ came up with a good analogy to relate the two techniques. Chang describes the difference between MSI and HSI to be like the differences between real analysis and complex analysis in mathematics. The intuitive interpretation that multispectral imaging is a special case of hyperspectral imaging is as incorrect as thinking that real analysis is a special case of complex analysis.

Because of the increased number of spectral bands, an HSI pixel contains more data than a multispectral imaging pixel. Multispectral pixels must rely on surrounding image pixels to provide spatial correlation and information to help make up for the limited spectral information, hence, why multispectral image processing techniques are spatial domain based. HSI pixels can contain information from targets that are subpixel size and, therefore, cannot be identified by visual inspection. Subpixel targets may not be detectable using the spatial domain-based image processing techniques developed for multispectral imaging. For this reason, hyperspectral image processing techniques generally involve target-based detection instead. Examples of hyperspectral targets may include objects like drug/smuggler trafficking, military vehicles, or in this case FAC/FIACs.

Polarization is an important physical quantity that describes how light interacts with an object during reflection, scattering, and transmission. A full description of polarized light can be seen in

equation 1.

$$S = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} I_0 + I_{90} \\ I_0 - I_{90} \\ I_{45} - I_{135} \\ I_{LHC} - I_{RHC} \end{bmatrix} \quad (1)$$

where I_0 , I_{45} , I_{90} , I_{135} are the intensity values of the linear polarization components at the angles denoted in the subscripts and I_{LHC} and I_{RHC} are the intensity values of left hand and right-hand circular polarization components. The S_0 and S_3 components of the Stokes vector are independent of the coordinate system, but the S_1 and S_2 components are dependent on the observation plane orientation. This polarization information can be measured using rotating polarizers, beam splitters, or focal plane array (FPA) division. For this project, beam splitters and FPA division is being investigated and will be discussed in more detail in section 3.2. It is necessary to acquire multiple images in different polarization angles to improve the object discrimination capabilities.

Polarization imaging techniques have been shown to improve signal-to-noise ratio and image contrast when compared to conventional imaging methods.¹⁵ The addition of polarization measurements can provide complementary information about how the imaged object is reflecting light. The four polarization components I_0 , I_{45} , I_{90} , I_{135} , provide information about the material properties and orientation of the object. The relationship between reflection coefficient values and incidence angle can be shown for 0- and 90-degree linear polarizations. This relationship is determined by equations 2 and 3. Where n_1 is the refractive index of the incident medium (marine air, in this case), n_2 is the transmissive medium (sea water), θ_1 is the angle of incidence, and θ_2 is the angle inside the second medium. The calculated reflection coefficients are shown in figure 2. Figure 2 shows that significant differences exist between the reflection coefficients of the two polarizations.

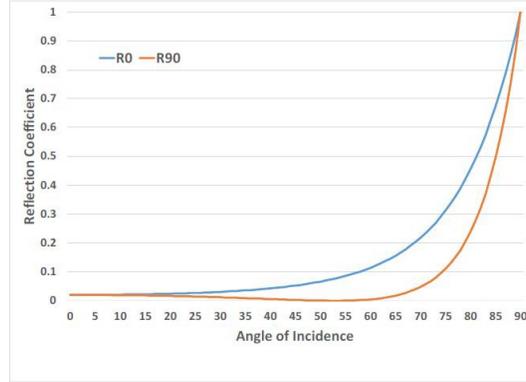


Fig 2 Reflection coefficient versus angle of incidence (from surface normal) for an air-water interface

The expectation is that capturing images in multiple polarization angles will improve detection and tracker performance. Multiple polarization images will be used in the tracking algorithm for motion detection and spatial analysis. A more in-depth description of this process can be found in section 4.4.

$$R_0 = \left| \frac{n_1 \cos(\theta_1) - n_2 \cos(\theta_2)}{n_1 \cos(\theta_1) + n_2 \cos(\theta_2)} \right|^2 \quad (2)$$

$$R_{90} = \left| \frac{n_1 \cos(\theta_2) - n_2 \cos(\theta_1)}{n_1 \cos(\theta_2) + n_2 \cos(\theta_1)} \right|^2 \quad (3)$$

3.2 Statements on Optical Design

The goal of the imaging system is to be able to detect small boats with a high degree of confidence at a distance of at least 10 km in noisy marine environment.¹⁶ The specific requirements of the imaging system can be seen below.

1. Spectral Response: Visible to Long Wave Infrared (LWIR)

- Visible: 0.4 - 0.7 μm

- Near Infrared (NIR): 0.7 - 0.9 μm
- Short Wave Infrared (SWIR): 0.9 - 1.7 μm
- Mid Wave Infrared (MWIR): 3.0 - 5.0 μm
- Long Wave Infrared (LWIR): 8.0 - 14 μm

2. Ability to switch between at least three field of views (FOVs);

- Narrow: Less than or equal to 3 degrees x 3 degrees
- Medium: 8 degrees x 8 degrees
- Wide: 15 degrees x 15 degrees

3. Boresighted and properly aligned sensors and FOVs

The three field of view (FOV) requirements were met by calculating the effective focal length of the lens used to image the object onto the detectors. This lens will need a variable focal length to be capable of switching between three FOVs. This calculation and a depiction of the problem can be seen in equation 5 and figure 3, respectively. To achieve the three required system FOVs of 3x3 degrees, 8x8 degrees, and 15x15 degrees, the effective focal length of the focusing lenses must be range from 396 mm to 78 mm. The system FOVs can be related to the minimum small boat detection criterion found in Krapels and Drigger's paper⁷ described in section 2.1. The reported spatial frequencies that resulted in a 50% detection rate were 14.0 cycles on target in the visible and 10.6 cycles for MWIR in the daytime. Using equation 6, we can calculate the minimum detector FOV to verify that the chosen detector size will be sufficient.

$$FOV = 2 \tan^{-1} \left(\frac{H}{2f_{eff}} \right) \quad (4)$$

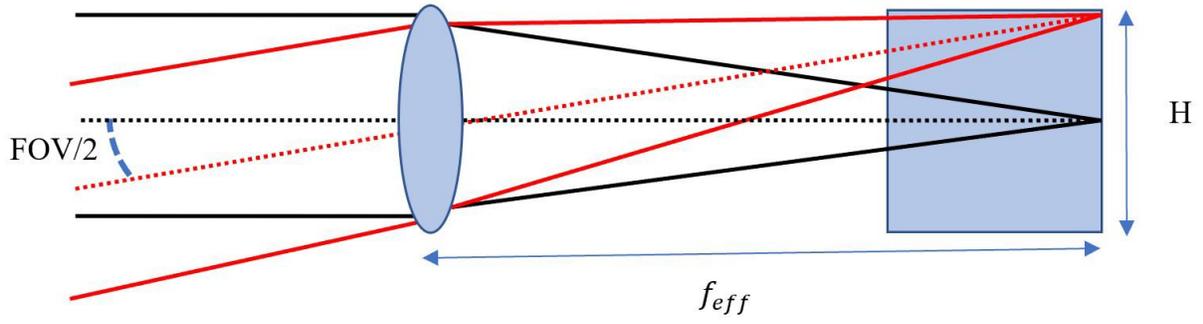


Fig 3 Block diagram of focusing lens and detector

$$f_{eff} = \frac{0.5H}{\tan(FOV/2)} \quad (5)$$

$$\Delta k = \frac{1}{FOV} \quad (6)$$

$$\Delta \omega = \frac{1}{k_{WFOV}}; \Delta \nu = \frac{1}{k_{VFOV}} \quad (7)$$

The use of multiple bands is necessary to improve imaging performance as imaging in littoral waters can prove to be difficult. The presence of aerosols in the water/air boundary layer,¹⁷ changing absorption of water as a function of wavelength,¹⁸ and the presence of marine life that may emit radiation in the bands of interest complicate the imaging process. As stated in section 2.1, studies have been done to show that the use of multiple wavelength bands helps improve imaging performance in a marine environment. Shaw and Burke,¹⁹ discuss how the selection of useful bands is dependent on the scattering and fluctuations of the atmospheric transition in the environment of

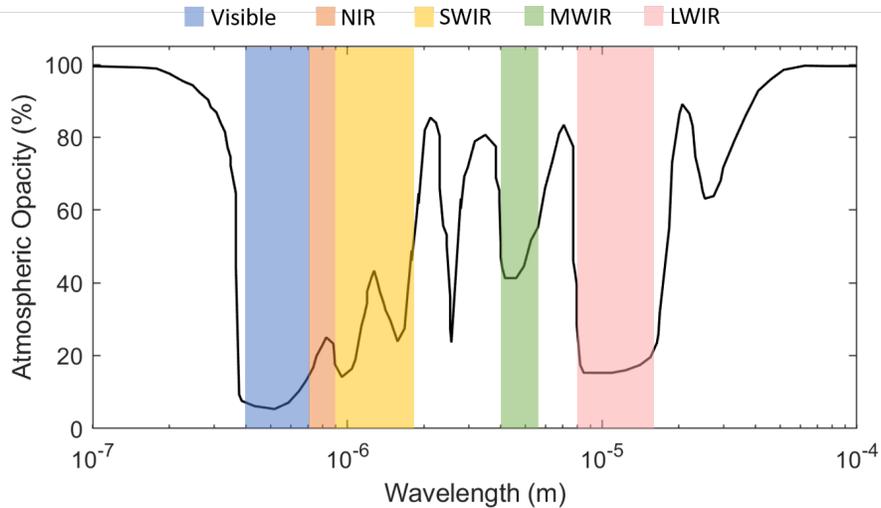


Fig 4 Plot of Atmospheric Opacity (%) versus Wavelength (m) with system spectral regions highlighted

interest.

The spectral bands chosen for this project are shown in list in section 3.2. These spectral bands were chosen by the Navy based off the ability to transmit through the atmosphere. Figure 4 shows the opacity of the atmosphere as a function of wavelength. The region of this plot that is of interest ranges from roughly 0.5 μm to 14 μm ; the spectral bands being used are highlighted. Each spectral band offers different information that can be utilized in small boat detection. The visible spectrum can be used to filter out clouds, smoke, and haze (blue band) and capture ‘true-color’ images, the NIR and SWIR bands are strongly absorbed by water creating a contrast difference between the boat and the ocean surface, and thermal IR (MW and LWIR) are used to detect heat signatures coming from the boats. It is clear, from figure 4, that these spectral bands have experience high transmission through the atmosphere.

As mentioned in section 3.1, light can be separated into its polarization components using rotating polarizers, polarized beam splitters, or FPA division. Issues arise with each of these methods, as with any design choices to be made in optics. The rotating polarizer may not move fast enough for proper data acquisition and result in target movement between polarization images ; beam split-

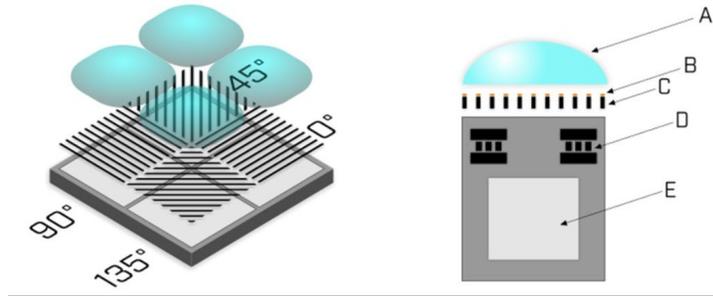


Fig 5 Plan view (left) and cross sectional view (right) on Sony’s IMX250MZR sensor. Each pixel’s polarizing filter (C) is coated with an anit-reflective layer (B) and is positioned between the microlens (A) and the light sensitive photodiode (E). Image from Sony Corporation²

ters require the use of multiple detectors and require careful alignment and matched FPAs; FPA division allows simultaneous collection and simplified alignment at the cost of image resolution. It was determined that the FPA division method will be the most effective way to meet the specified requirements. FPA division was chosen for the visible and NIR bands to reduce system space by reducing the number of components by six sensors and four beam splitters (when compared to division of amplitude method), while still providing sufficient resolution to effectively perform target detection. Sony Corporation (Japan) released a CMOS sensor with integrated polarizers.² Each individual pixel has its own polarizing filter, at polarization angles of 0, 45, 90, and 135 degrees. The filters are arranged in two-pixel blocks as shown in figure 5.

Sony announced a 12.3MP version of this sensor that will allow 1K x 1K imaging of the first three Stokes components (S_0 , S_1 , S_2) in a single camera. The CMOS camera can be used for polarization separation in the visible and NIR bands. As of late, a camera with integrated polarizers has not been found to work in the SWIR, MWIR, and LWIR bands. For these spectral bands, polarized beam splitters and multiple (4) cameras will need to be used. The preliminary design for this configuration can be seen in figure 6. Figure 6 shows a ‘flattened’ view of the four-camera measurement system for a single spectral band. The actual configuration will have fewer

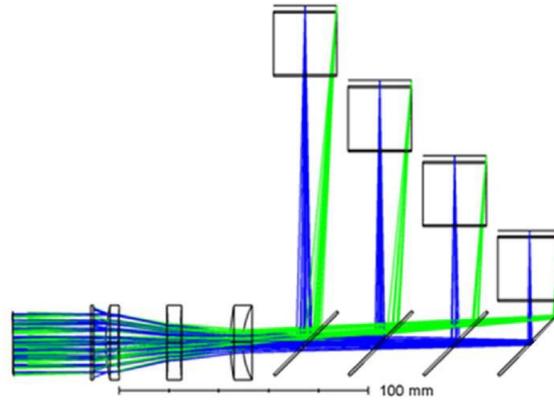


Fig 6 Preliminary Zemax optical model of the five band imaging system

beam splitters and be folded in three dimensions to minimize the overall volume.

4 Automatic Target Recognition (ATR)

4.1 ATR Background

Automatic target recognition (ATR) is the ability of an algorithm to recognize objects based on data obtained from sensors. ATR minimizes the need for a human operator to manually select the desired targets in image data. Removing manual selection from target recognition process allows for real-time tracking. There are many methods that can be used for target recognition, some can be seen in section 2.2. The following sections will also go into more detail on two more detection methods that are being used to detect FAC/FAICs in the ocean. These detection methods are Kalman filtering and Reed – Xiaoli anomaly detection.

4.2 Data Fusion

Initial processing must take place to align and normalize data collected across multiple sensors with different pixel counts, densities, and sensitivities. These basic operations can be seen in the left-hand column in figure 7. After proper alignment and normalization, the data can be ‘fused’

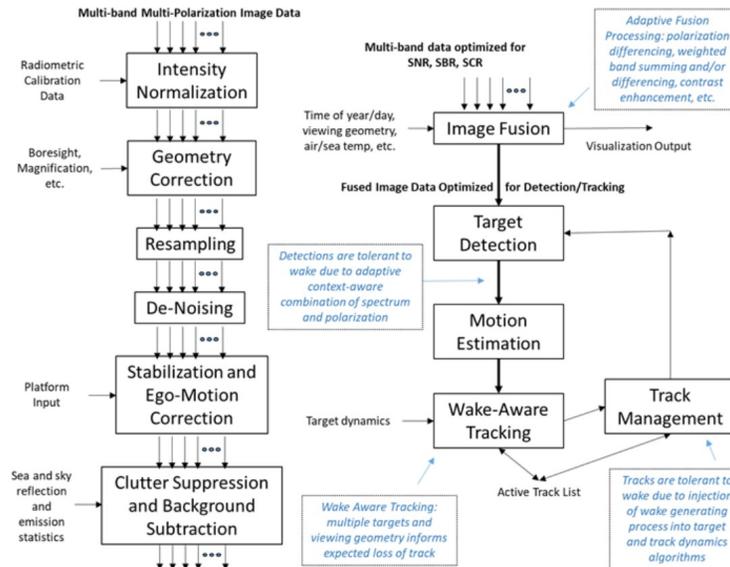


Fig 7 Preliminary Data Fusion and Analysis Model (courtesy of Dr. Mark Neifeld)

for single-frame spatial analysis. Fusion of the spectral bands and polarization channels provides the greatest contrast for shape detection. Each of the polarization channels are analyzed separately before being combined to determine how tracking performance is affected. The right-hand column of figure 7 shows how the target detection loop will occur and introduces how the target detection method will be tolerant to the boat wake due to the adaptive data fusion process that was implemented. The wake-aware tracking accounts for multiple targets and viewing geometries to determine points of expected loss of tracking.

4.3 Kalman Filtering

A Kalman filter is a type of state observer that is used for stochastic systems (rather than deterministic systems). Kalman filtering is a two-step process that involves a prediction of some measurement and updating predictions based on said measurements. Kalman filtering is a recursive algorithm so these updated models are then used for the next states predictions and the algorithm repeats itself. The prediction of states and the measurements are represented as probability density

functions (PDF) with normal distributions and specified means and variances. A Kalman filter optimizes the extent to which the predicted estimate and the measurement are to be trusted in the updated model. The extent of the confidence depends on the variance of each distribution, smaller variance results in a thinner PDF and a higher probability of landing at the mean of the measurement. Once the weights of the predicted state estimate and the measurement state are determined, the two PDFs are multiplied together to create an updated PDF, also with a normal distribution, with a mean somewhere in between the previous values and a smaller variance. Kalman filters are often the best selection for video-based tracking because they provide a linear solution for discrete, (frame rate) stochastic, dynamic processes. The time evolved Kalman filter can be described by discrete linear equations (eqns. 8 and 9).

$$x(t+1) = A * x(t) + B * u(t) + w(t) \quad (8)$$

$$z(t) = C * x(t) + D * u(t) + v(t) \quad (9)$$

Where x is the process state vector, u is the control vector, z is the measurement vector obtained by a tracking algorithm, A is the state transition matrix, B is the state control matrix, C is the observation matrix, D is the measurement control matrix, w is the noise associated with the state, v is the noise associated with the measurement, and Q and R are the diagonal covariance matrices. The process state vector x contains variables related to the object translation, scale and orientation, and its velocity and acceleration.²⁰

As stated previously, Kalman filtering is a recursive algorithm that consists of two steps: state

prediction and measurement update. The state prediction phase (equations 10 and 11) predicts the state vector at a time x , and the error matrix P . It is important to remember that the state vector and error matrix are conditional probability predictions given the observations obtained at $t-1$. Equations 10 and 11 show how the state predictor phase is determined.

$$x(t|t-1) = A * x(t-1|t-1) \quad (10)$$

$$P(t|t-1) = A * P(t-1|t-1) * A + Q \quad (11)$$

where $x(t)$ is the state at frame t , $x(t|t-1)$ and $x(t-1|t-1)$ are a priori and a posteriori estimation of the state vector, $P(t|t-1)$ and $P(t-1|t-1)$ are the a priori and a posteriori estimation of the error matrix, and E is the expected value. The measurement update phase corrects the state vector $x(t|t-1)$ value and the error matrix $P(t|t-1)$ values by accounting for the $z(t)$ measurement obtained by the tracker at each frame as shown in equations 12 - 14, where $K(t)$ is the Kalman gain at frame t .

$$K(t) = P(t|t-1) * C^t * (C * P(t|t-1) * C^T + R)^{-1} \quad (12)$$

$$x(t|t) = x(t|t-1) + K(t) * (z(t) - C * x(t|t-1)) \quad (13)$$

$$P(t|t) = P(t|t-1) - K(t) * C * P(t|t-1) \quad (14)$$

Kalman filtering has been the foundation for several marine tracking algorithms including succes-

sive clustering,²¹ optical flow,¹⁰ mean shift tracking,¹¹ histogram matching,²² and active contour tracking.²³ Each of these listed methods have their strengths and weaknesses based on single band imaging typically used in marine tracking.

It is common for FAC to engage in tactics that intentionally try to obscure the true number and/or position of boats. The obscuration of boats could lead to a collapse of the Kalman filter update phase as no new measurements on the ‘hidden’ boats can be made. The recursive algorithm would then have to be reset resulting in significant delays in tracking. It is hypothesized in this project that the multiple wavelength bands and polarization inclusion will help prevent the breakdown of tracking algorithms because of boat obscuration.

4.4 Reed – Xiaoli Anomaly Detection

Moving targets can be located by detecting statistical deviations in pixels across image frames (i.e., pixel changes over time). Here, it is assumed that the variations in the value of a background pixel are due to temporal noise. The temporal noise is assumed to be additive white Gaussian noise (AWGN) with a normal distribution of background pixels. As an object moves across the window of pixels being observed, the quantity shown in equation 15 will be large.

$$\log[p_b(v_k)] \approx -\frac{1}{2}(v_k - \mu)^T \Sigma^{-1}(v_k - \mu) \quad (15)$$

$$p_b(v_k) = N(\mu, \Sigma) \quad (16)$$

This equation is the characteristic equation of a Reed – Xiaoli anomaly detector (RXD) where $p_b(v_k)$ is the probability that a pixel will be a part of the background, v_k is a vector with the

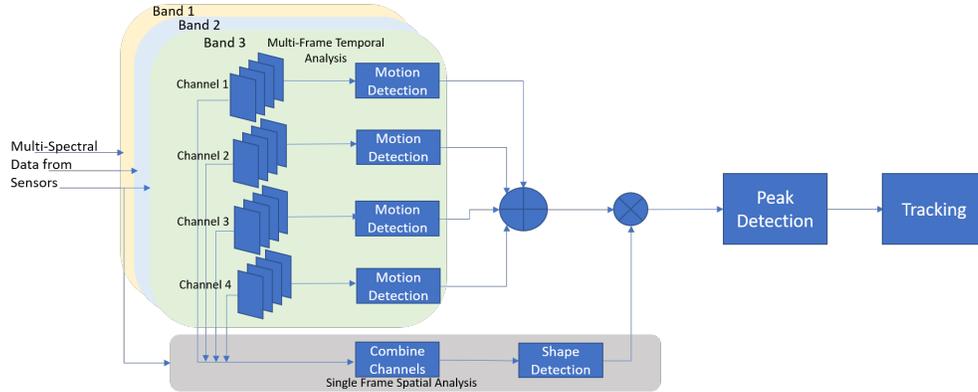


Fig 8 Tracking algorithm system architecture (credit: Dr. Abhijit Mahalanobis)

values observed at the k -th pixel over several frames of data, μ is the mean value of the AWGN distribution, and Σ is the covariance of the AWGN distribution. The log quantity reaches its maximum when the pixel in question has a low probability of being a background pixel. A pixel has a low probability of being a background pixel when the variations in pixel intensity deviate from the AWGN distribution. The temporal analysis is done by arranging a pipeline of multiple frames for each polarization and spectral band. The RXD is then applied to detect the object motion over the set of frames. A depiction of this system architecture is shown in figure 8. An RXD anomaly score is then given to each of the channels and the channels are added to maximize probability of detection. Detection results will be discussed in section 4.5. It is also noteworthy to mention that single frame spatial analysis is done with the combined channels to perform shape detection. This process supplements the RXD motion detection and allows for increased detection probability.

4.5 Detection Results

Receiver operating characteristic (ROC) curves were generated for each of the separate channels. Figure 9 shows the ROC curves resulting from the RXD algorithm. It is clear the two of the polarization channels (channels 2 and 3) outperform the other two channels. A frame from each of the polarization channels is shown in figure 10. Channels 2 and 3 show a higher contrast between

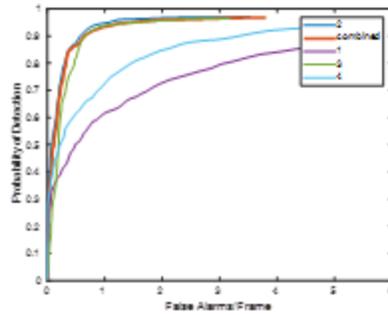


Fig 9 ROC curves from RXD results (credit: Dr. Abhijit Mahalanobis)

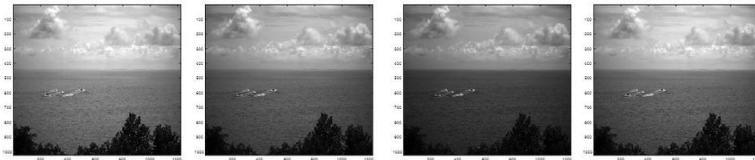


Fig 10 Single Frames from each polarization channel (1-4) and same spectral band

the ocean and the boats than channels 1 and 4, so it can be expected that the tracker will have an easier time detecting the boats in channels 2 and 3. Figure 9 also shows an ROC curve generated for the combined channels. The ‘combined’ curve is obtained by adding the RXD scores from each of the channels, hence why it is pulled up to the highest performing polarization channel. Currently, it is difficult to determine whether combining each of the polarization channels is beneficial in improving tracker performance.

Motion detection was performed in two scenarios: with and without single frame spatial analysis. When tracking was done with detection based only on RXD scores, the average track duration averaged 35%. This means that 35% of the total boat path was recognized by the tracking algorithm. When tracking was done with detection based on the product of RXD scores and spatial analysis scores, the average track duration increased to 55%. While it is not clear whether adding the polarization channels improves tracker performance, it is clear that motion analysis weighted with spatial analysis scores do improve tracker performance.

5 Verification via Ground Truthing

5.1 Ground Truth Data Background

Ground truth data contains a set of images, a set of labels on the images, and a model of defining the detected objects. This model should specify the count, location, and other key features about the detected objects. Carefully selected ground truth data provides a robust method of measuring a tracker algorithm's precision and accuracy. Being able to properly understand performance and accuracy of tracking algorithms is the primary purpose for establishing ground truth data. Ground truth data can vary for different tasks. Krig²⁴ gives the examples of 3D image reconstruction and face recognition, one must recognize the different attributes of the ground truth data for each task. For face recognition, some of the ground truth data may require segmentation and labeling to define face location and orientation, size of face, and physical characteristics such as emotion, gender, and age. On the other hand, 3D reconstruction may need the raw pixels in the images and a reference mesh or point cloud for ground truthing.

There are several categories that describe ground truth datasets:²⁴

- Synthetic produced: images are generated from computer models or renderings.
- Real produced: a video or image sequence is designed and produced.
- Real selected: real images are selected from existing sources.
- Machine-automated annotation: feature analysis and learning method are used to extract features from the data.
- Human annotated: an expert defines the location of features and objects.

- Combined: any mixture of the above.

For this project, a ‘real selected’ ground truth dataset was created and used for validation of the tracking algorithm. This biggest challenge with using a real selected dataset is that a person must spend the time to create that dataset. This process is very tedious and can take many hours as, in this case, there are hundreds of frames and multiple objects to be selected on each frame. Since the collected data and tracking algorithm were developed specifically for this project, there are no previously existing ground truth datasets that can be used. A new ground truth dataset had to be established to effectively measure the performance of the tracking algorithm. A tool was developed using MATLAB App Designer™ to establish the ground truth dataset used in this project.

The established ground truth data is used to quantitatively measure the accuracy of the developed tracking algorithm. Quantitative accuracy measurements are accomplished by comparing the pixel coordinates for each object, provided by the tracking algorithm, to the pixel coordinates specified in the ground truth data. This analysis can be used to generate an error matrix and pinpoint when/where location errors are most often occurring. Along with monitoring location errors, ground truth data can be used to determine false positive and false negative occurrences. False positives occur when the tracking algorithm detects an object that is not actually an object of interest and false negatives occur when the tracking algorithm does not detect an object of interest at all. Properly established ground truth data will contain information from each and every object of interest in the video in question, allowing us to determine when/where errors of commission and omission occur.

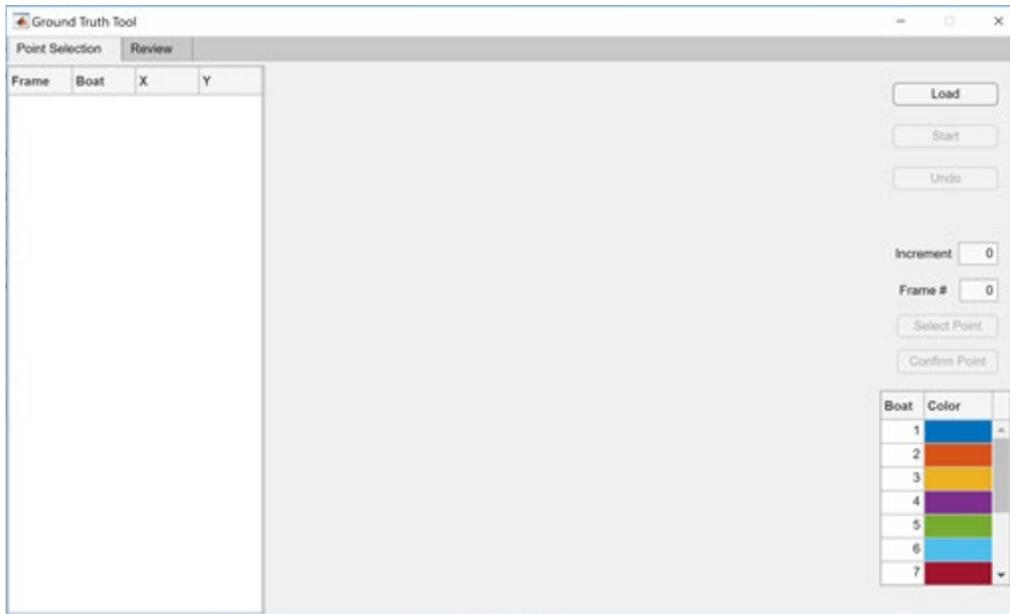


Fig 11 Screenshot of the GUI from the developed Ground Truth App

5.2 *Developed Ground Truth Tool*

It was necessary to develop a tool that can be made easily accessible to a team and provide an intuitive user interface to make the ground truthing process as efficient as possible. To accomplish this, a ground truthing app was created in MATLAB App Designer™. A screenshot of the graphical user interface (GUI) can be seen in figure 11. There are buttons to load in the video files in need of ground truth datasets, start the selection process, undo a selection, save the corresponding data, and buttons used to select and confirm object selection. Once object selections are confirmed, the frame number the object is located on, the object number, and the X and Y coordinates of the object (in pixel values), are presented in the table to the left of the image.

Figure 12 shows the different tabs available in the GUI. There is a “Point Selection” tab that is used for the manual object selection and a “Review” tab that is used to view the selected points plotted on top of a frame from the selected video. The “Review” tab is used to ensure that all selected points came out as expected and there are no obvious errors. The app also has the option

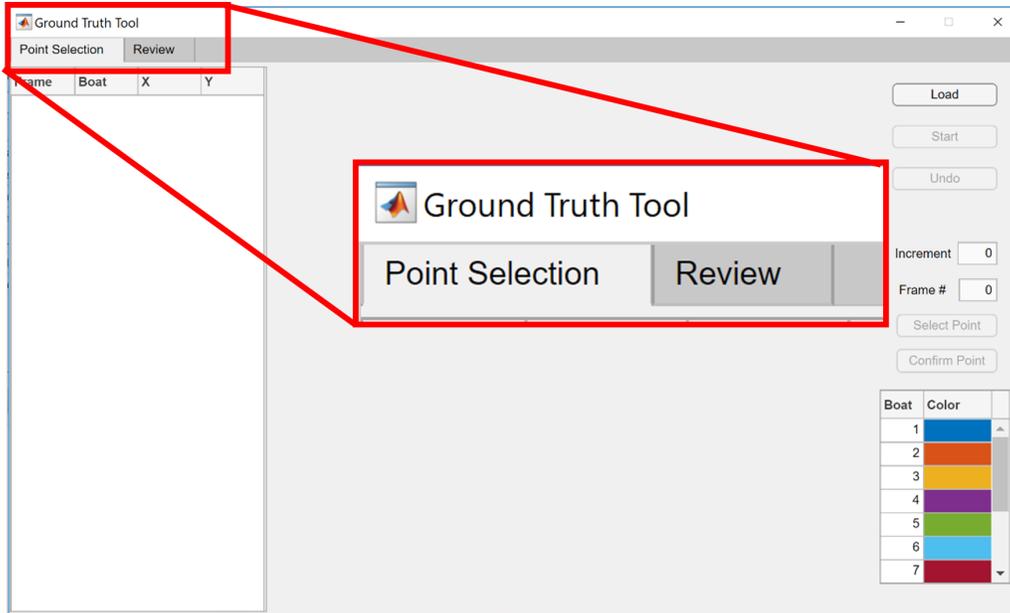


Fig 12 Screenshot of the GUI from the developed Ground Truth App with Tab selection highlighted

to skip a certain number of frames that is deemed appropriate by the user. This can be used, for example, when the objects do not show obvious movement in consecutive frames. This feature helps eliminate redundant data and speeds up the ground truth process. In case the user defines a frame increment that is too large the app has a built-in linear interpolation function that will estimate the position of each object on the frames that were skipped. The interpolated points will be plotted as stars, while the selected ground truth points will be plotted as filled in circles on the “Review” plot. The hope is that the “Review” plot will provide the user with the information to decide if a finer frame increment is necessary or not. A screenshot of the “Review” plot is shown in figure 13.

It is important to note that the data presented in the table in the “Point Selection” tab is not the same as the data presented in the “Review” tab. The data presented in the “Point Selection” tab consists only of the manually selected data points, while the data presented in the “Review” tab consists of the manually selected data points and each of the interpolated points. To differentiate

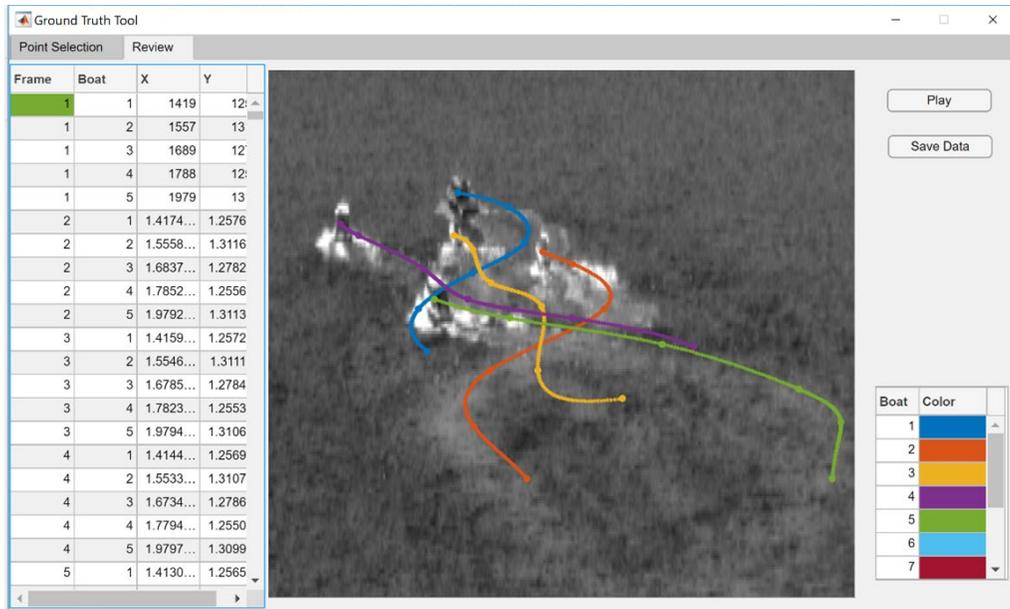


Fig 13 Screenshot of the "Review" tab after point selection is finished

the points that were manually selected from the points that were interpolated, the cells that contain frames where selection occurred are highlighted green. A screenshot of the two tables with example data is shown in figure 14.

The point selection process involves two button presses: 1. To initiate the point selection (presumably after all necessary zoom/pan have occurred to accurately select boat position) and 2. To confirm that the selected point is accurate and can be entered into the table. If the initial point selected is an error, a new point can be selected (before confirmation of the first) to replace the prior selection. Selection can take place as many times as needed before confirming the proper data point. If the user finds themselves in a position where they have confirmed a data point that they did not wish to confirm, there is an "Undo" button that will restart the selection process for the current frame. The "Undo" button will clear all the selected/confirmed points from the data table for that frame and the user will have to start again. It is hoped that the user will rarely have to use this function as they are given the option to re-select an infinite number of times before confirming the selected point.

Point Selection		Review	
Frame	Boat	X	Y
1	1	1419	12
1	2	1557	13
1	3	1689	12
1	4	1788	12
1	5	1979	13
52	1	1407	12
52	2	1484	12
52	3	1572	12
52	4	1619	12
52	5	1991	12
103	1	1481	12
103	2	1506	12
103	3	1577	12
103	4	1475	12
103	5	1933	12
154	1	1553	12
154	2	1664	12
154	3	1507	12
154	4	1413	12
154	5	1744	12
205	1	1531	11

Point Selection		Review	
Frame	Boat	X	Y
1	1	1419	12
1	2	1557	13
1	3	1689	12
1	4	1788	12
1	5	1979	13
2	1	1.4174...	1.2576
2	2	1.5558...	1.3116
2	3	1.6837...	1.2782
2	4	1.7852...	1.2556
2	5	1.9792...	1.3113
3	1	1.4159...	1.2572
3	2	1.5546...	1.3111
3	3	1.6785...	1.2784
3	4	1.7823...	1.2553
3	5	1.9794...	1.3106
4	1	1.4144...	1.2569
4	2	1.5533...	1.3107
4	3	1.6734...	1.2786
4	4	1.7794...	1.2550
4	5	1.9797...	1.3099
5	1	1.4130...	1.2565

Fig 14 Screenshot of the tables on the "Point Selection" tab (left) and "Review" tab (right)

To provide further guidance in point selection, the location of each object from the previous frame is plotted on the current frame. Each object is assigned a number (based on click order) and a corresponding color that can be found in the tables in the lower right-hand corner of the GUI. Plotting the previous location of each object is expected to provide a trajectory for each of the objects and assist in keeping the initial click order. Each of the datasets containing just the selected points and both the selected and interpolated points can be saved as ASCII text files. These files can be loaded into Excel™ or MATLAB™ for processing.

6 Conclusions

A general layout of the multispectral polarimetric imaging system has been presented along with a high-level description of the accompanying tracking algorithm. Within the optical design, the importance of the chosen spectral bands, focal plane resolution, and polarization was described.

The spectral bands chosen for this prototype: visible (0.4 – 0.7 μm), NIR (0.7 – 0.9 μm), SWIR (0.9 – 1.7 μm), MWIR (3.0 – 5.0 μm), and LWIR (8.0 – 12.0 μm). These bands show high atmospheric transmission and provide necessary contrast data for target discrimination and tracking. It was also noted that the minimum pixel count for the focal plane array must be at least 1K x 1K for accurate FAC detection at 10 km at the widest FOV (15 degrees). Finally, the importance of polarization in the different bands to increase sensitivity in boat wakes and features was shown. The tracking algorithm was developed using Kalman filtering and (RXD) with verification of results produced via the developed ground truth tool. Detection and tracking results were also reported, and it was found that including single-frame spatial analysis along with motion detection improved overall tracking performance (from 35% track duration to 55%). The developed ground truth tool provides an effective method of establishing ground truth data for the collected images. The established ground truth data will be used to quantitatively analyze the performance of the tracking algorithm and help determine where tracking issues may occur. This tool achieves the goal of being easily accessible by being a stand-alone desktop app that can be downloaded and ran on any computer.

APPENDIX A – TABLE OF ACRONYMS

Acronym/Initialism	Expression
FAC/FAIC	Fast Attack Craft/Fast Attack Inshore Craft
F2T2EA	Find Fix Track Target Engage Assess
EO/ IR	electro optical/ infrared
NIR	near infrared
SWIR	shortwave infrared
MWIR	midwave infrared
LWIR	long wave infrared
PBS	polarizing beam splitter
DR	detection rate
FAR	false alarm rate
ATR	automatic target recognition
CNN	convolutional neural network
DCNN	deep convolutional neural network
FLIR	forward looking infrared
MSI	multi spectral imaging
HSI	hyperspectral imaging
GSD	ground sampling distance
FPA	focal plane array
FOV	field of view
CMOS	complementary metal oxide semiconductor
PDF	probability density function
AWGN	additive white gaussian noise
RXD	Reed – Xiaoli anomaly detection
ROC	receiver operating characteristic
GUI	graphical user interface

References

- 1 F. Snik, J. Craven-Jones, M. Escuti, *et al.*, “An overview of polarimetric sensing techniques and technology with applications to different research fields,” in *Polarization: Measurement, Analysis, and Remote Sensing XI*, D. B. Chenault and D. H. Goldstein, Eds., **9099**, 48 – 67, International Society for Optics and Photonics, SPIE (2014).
- 2 https://www.sony-semicon.co.jp/products/common/pdf/IMX250_264_253MZR_MYR_Flyer_en.pdf
- 3 S. LaGrone, “U.s. navy surveillance ship harassed by iranian attack boat,” (2017).
- 4 X. Rice, “Us navy captures somali pirates’ mother ship,” (2010). Section: World news.
- 5 R. Hassig and H.-S. Lee, “North korea: A strange socialist fortress,” *Journal of Asian Studies - J ASIAN STUD* **61** (2002).
- 6 J. R. Buss and G. R. Ax, “Fleet protection using a small uav based ir sensor,” 17 (2005).
- 7 K. Krapels, R. G. Driggers, D. Deaver, *et al.*, “Midwave infrared and visible sensor performance modeling: small craft identification discrimination criteria for maritime security,” *Appl. Opt.* **46**, 7345–7353 (2007).
- 8 A. W. Cooper, W. J. Lentz, P. L. Walker, *et al.*, “Infrared polarization measurements of ship signatures and background contrast,” in *Characterization and Propagation of Sources and Backgrounds*, W. R. Watkins and D. Clement, Eds., **2223**, 300 – 309, International Society for Optics and Photonics, SPIE (1994).
- 9 D. Bloisi, L. Iocchi, M. Fiorini, *et al.*, “Automatic maritime surveillance with visual target detection,” (2011).
- 10 D. Bloisi and L. Iocchi, “Argos - a video surveillance system for boat traffic monitoring in venice.,” *IJPRAI* **23**, 1477–1502 (2009).
- 11 D. Comaniciu, V. Ramesh, and P. Meer, “Kernel-based object tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(5), 564–577 (2003).
- 12 N. Nasrabadi, “Deeptarget: An automatic target recognition using deep convolutional neural networks,” *IEEE Transactions on Aerospace and Electronic Systems* **PP**, 1–1 (2019).
- 13 P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, **1**, I–I (2001).
- 14 C. Chang, *Hyperspectral Data Exploitation: Theory and Applications*, Wiley (2007).
- 15 J. S. Tyo, D. L. Goldstein, D. B. Chenault, *et al.*, “Review of passive imaging polarimetry for remote sensing applications,” *Appl. Opt.* **45**, 5453–5469 (2006).
- 16 L. LaComb, M. Neifeld, and A. Amit, “Multispectral/hyperspectral imaging system for small boat detection under wake clusters,” (2019).
- 17 Z. Kozarac, D. Risović, S. Frka, *et al.*, “Reflection of light from the air/water interface covered with sea-surface microlayers,” *Marine Chemistry* **96**, 99–113 (2005).
- 18 G. M. Hale and M. R. Querry, “Optical constants of water in the 200-nm to 200- μ m wavelength region,” *Appl. Opt.* **12**, 555–563 (1973).
- 19 G. Shaw and H.-h. Burke, “Spectral imaging for remote sensing,” *Lincoln Laboratory Journal* **14** (2003).
- 20 R. Moreira, N. Ebecken, A. S. Alves, *et al.*, “A survey on video detection and tracking of maritime vessels,” *International Journal of Research and Reviews in Applied Sciences* **20** (2014).

- 21 S. Fefilatyev, D. Goldgof, and C. Lembke, “Tracking ships from fast moving camera through image registration,” in *2010 20th International Conference on Pattern Recognition*, 3500–3503 (2010).
- 22 J. Ning, L. Zhang, D. Zhang, *et al.*, “Joint registration and active contour segmentation for object tracking,” *IEEE Transactions on Circuits and Systems for Video Technology* **23**(9), 1589–1597 (2013).
- 23 R. Goldenberg, R. Kimmel, E. Rivlin, *et al.*, “Fast geodesic active contours,” **1682** (1999).
- 24 S. Krig, “Ground Truth Data, Content, Metrics, and Analysis,” in *Computer Vision Metrics: Textbook Edition*, 247–271, Springer International Publishing, Cham (2016).