

University of Memphis

University of Memphis Digital Commons

Electronic Theses and Dissertations

10-13-2022

An Impulse Detection Methodology and System with Emphasis on Weapon Fire Detection

Jeremy Benjamin Brown

Follow this and additional works at: <https://digitalcommons.memphis.edu/etd>

Recommended Citation

Brown, Jeremy Benjamin, "An Impulse Detection Methodology and System with Emphasis on Weapon Fire Detection" (2022). *Electronic Theses and Dissertations*. 3201.

<https://digitalcommons.memphis.edu/etd/3201>

This Dissertation is brought to you for free and open access by University of Memphis Digital Commons. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of University of Memphis Digital Commons. For more information, please contact khhgerty@memphis.edu.

AN IMPULSE DETECTION METHODOLOGY AND SYSTEM WITH EMPHASIS ON
WEAPON FIRE DETECTION

by

Jeremy Benjamin Brown

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Electrical and Computer Engineering

The University of Memphis

December 2022

Distribution Statement A: Approved for public release. Distribution is unlimited.

DEDICATION

This dissertation is dedicated to mother, Carolyn Brown, my father, Percy Ken Brown, and those that came before, for never giving up on me, and for creating a path forward, that I might follow.

ACKNOWLEDGEMENTS

The Lord journeys before me, and walks beside me, a light in dark places, blessing me before I begin, strengthening me when in need, and encouraging me that I might not be afraid. The Lord shields me, even from myself, shows me the way, and, what is to come. He puts thoughts to mind, that I might take to action. He builds me up and prepares me, so that, I can stand in front of others. He softens my heart, so I, might hear their voices. He gives me understanding, so that I can see the soul. He places me, where, I am to be. He makes me ready, so, that I am able, to do the work, He, lays before me. I lay before the Lord. Thank you Lord.

I would like to thank my parents for their continued support. My mother was part of the integrating class at Benjamin Franklin High School in New Orleans, Louisiana. I later attended this same high school, some 30 years later, receiving a premier education at one of the top high schools in Louisiana. It was her courage those many years ago, that made today possible. My parents always made sure I had what I needed to succeed. Though we were not wealthy, we had an abundance of love and education. Thank you.

I would also like to thank my major advisor, Dr. Eddie Jacobs, who has also “stuck with me,” or has been “stuck with me,” for quite some time. The direction of my professional career, opportunities, and professional expertise, would not have been possible without Dr. Jacobs. Additionally, I am so grateful to Dr. Chris Howell, the first African-American PhD graduate at the University of Memphis School of Engineering, and Dr. Aaron Robinson, for taking the time and interest to convince me to attend graduate school, at which time, had not been a part of my imagination or reality. Each of them has encouraged and pushed me to this point.

Additionally, I would thank the University of Memphis faculty, classmates, colleagues, friends, family, and loved ones who have supported me over the years, and I hope that I can provide the same encouragement to others, that I was fortunate enough to receive.

ABSTRACT

Brown, Jeremy Benjamin. Ph. D. The University of Memphis. December 2020. An Impulse Detection Methodology and System with Emphasis on Weapon Fire Detection. Major Professor: Dr. Eddie L. Jacobs.

This dissertation proposes a methodology for detecting impulse signatures. An algorithm with specific emphasis on weapon fire detection is proposed. Multiple systems in which the detection algorithm can operate, are proposed. In order for detection systems to be used in practical application, they must have high detection performance, minimizing false alarms, be cost effective, and utilize available hardware. Most applications require real time processing and increased range performance, and some applications require detection from mobile platforms. This dissertation intends to provide a methodology for impulse detection, demonstrated for the specific application case of weapon fire detection, that is intended for real world application, taking into account acceptable algorithm performance, feasible system design, and practical implementation.

The proposed detection algorithm is implemented with multiple sensors, allowing spectral waveband versatility in system design. The proposed algorithm is also shown to operate at a variety of video frame rates, allowing for practical design using available common, commercial off the shelf hardware. Detection, false alarm, and classification performance are provided, given the use of different sensors and associated wavebands. The false alarms are further mitigated through use of an adaptive, multi-layer classification scheme, leading to potential on-the-move application. The algorithm is shown to work in real time. The proposed system, including algorithm and hardware, is provided.

Additional systems are proposed which attempt to complement the strengths and alleviate the weaknesses of the hardware and algorithm. Systems are proposed to mitigate saturation

clutter signals and increase detection of saturated targets through the use of position, navigation, and timing sensors, acoustic sensors, and imaging sensors. Furthermore, systems are provided which increase target detection and provide increased functionality, improving the cost effectiveness of the system.

The resulting algorithm is shown to enable detection of weapon fire targets, while minimizing false alarms, for real-world, fieldable applications. The work presented demonstrates the complexity of detection algorithm and system design for practical applications in complex environments and also emphasizes the complex interactions and considerations when designing a practical system, where system design is the intersection of algorithm performance and design, hardware performance and design, and size, weight, power, cost, and processing.

PREFACE

Two patents resulting from the research described in this dissertation are used within the manuscript of this dissertation. The patents were accepted and published by the United States Patent and Trademark Office (USPTO) as US 10,389,928 B2 “Weapon Fire Detection and Localization Algorithm for Electro-Optical Sensors” (8/20/2019), and US 10,209,343 B1 “Weapon Fire Detection and Localization System for Electro-Optical Sensors” (2/19/2019).

Chapter 4 uses patent submission from US 10,209,343 B1 “Weapon Fire Detection and Localization System for Electro-Optical Sensors” (2/19/2019). Chapters 2 and 3 further describe the methodology presented in Chapter 4. Chapters 2 and 3 are written for eventual journal publication. Chapter 4, additionally, uses patent submission from US 10,209,343 B1 “Weapon Fire Detection and Localization System for Electro-Optical Sensors” (2/19/2019). A third patent, titled “False Positive Determination of a Blast”, has also been filed to the USPTO under application number 15/998,345 (8/9/2018).

Additionally, the work defined in Chapters 2, 3, and 4, contributed to the creation of two journal papers. The journal papers, “Weapon Fire Detection: Clutter and Saturation Mitigation” and “Advancements in Weapon Fire Detection Methodologies” were submitted to the Journal of Department of Defense Research and Engineering (JDRE) (3/8/2021).

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
I. INTRODUCTION	1
Objective.....	1
Hypothesis	4
Novelty	4
Author Contributions	6
Background	11
Detection Theory	11
General Detection Systems.....	39
Muzzle Flash and Blast Phenomenology.....	40
Optical Detection Methods and Systems.....	47
Multi-Sensor Methodology and Systems	48
Acoustic Detection Phenomenology	49
Acoustic Detection Methods & Systems.....	54
Position, Navigation, and Timing.....	59
Alternate Applications.....	61
Organization of this Dissertation	63
II. DETECTION AND CLASSIFICATION METHODOLOGIES.....	64
Introduction.....	64
Problem Description	64
Methods	64
Pre-Detections	66
Detection.....	69
Classification	75
Results.....	76
Detection.....	76
Classification	82
Discussion	87
III. CLUTTER MITIGATION METHODOLOGIES.....	89

Introduction.....	89
Problem description	89
Methods	89
Signatures	89
Algorithm.....	90
Results.....	105
Detection.....	105
Classification	109
Real-Time Operation	115
Discussion	116
IV. DETECTION, CLASSIFICATION, AND LOCALIZATION SYSTEM AND METHODOLOGY	118
Introduction.....	118
Background	119
Problem Description	122
Methods	123
Discussion	141
V. CONCLUSIONS, CONTRIBUTIONS, AND FUTURE WORK.....	142
Conclusion.....	142
Contributions.....	142
Future Work	143
REFERENCES.....	145

LIST OF TABLES

Table	Page
Table II.1. Probability of detection for hostile fire events using data collected from MWIR-1, MWIR-2, and LWIR imagers.	76
Table II.2. On-the-move false alarm results for a cloudy day. (Top) False alarms after each layer of filtering. (Bottom) False alarms per clutter environment.	78
Table II.3. On-the-move false alarms across cloudy and sunny days per clutter environment. General trends in false alarm rates per spectral band and across environments can be observed.	78
Table II.4. Classification results of LDA, QDA, Tree, and SVM using leave-one-out cross validation.....	83
Table II.5. False alarm reduction from the detection stage, using an adaptive false alarm class approach, and a distance from hostile fire approach.....	87
Table III.1. Classifier performance at 120 Hz per band, using groundtruth data.	110
Table III.2. Classifier performance at 600 Hz, using groundtruth data.	110
Table III.3. False detections and false alarms per band as a function of weather condition for 120 Hz operation.....	113
Table III.4. False detections and false alarms per band as a function of scene motion for 120 Hz operation.	114
Table III.5. Confusion matrix for dual band LWIR/MWIR-1 at 120 Hz. Undetected hostile fire events are also included.	115

LIST OF FIGURES

Figure	Page
Figure I.1. Diagram depicting system design as the intersection of sensor design, algorithm design, and SWAP-C	4
Figure I.2 Signal detection theory definitions and assumptions. [1]	14
Figure I.3. Distributions with unequal standard deviations. [1]	15
Figure I.4. Contrast vs spatial frequency. Contrast is varied vertically. Spatial frequency is varied horizontally.	19
Figure I.5. Irregular shaped target mask	23
Figure I.6. Background mask corresponding to Figure I.5 using a process of perimeter thickening	24
Figure I.7. Target (grey) with bounding box background (blue) of similar proportion and equal size	25
Figure I.8. Target (grey) with bounding box background (blue), partially outside of the image border (red)	26
Figure I.9. Histogram of background with $\mu=128$ and $\sigma=5$. Target pixel (red) with distance of 30 from the background μ is distinct from the background.....	29
Figure I.10. Histogram of background with $\mu=128$ and $\sigma=20$. Target pixel (red) with distance of 30 from the background μ is not distinct from the background.	30
Figure I.11. Distribution of sample observations with two independent features	32
Figure I.12. Distribution of sample observations with two dependent features	32
Figure I.13. Shadowgraph of precursor flow. (1: blast wave; 2: Mach cone; 3: barrel shock; 4: spherical flow discontinuity; 5: gas-air interface) [17].....	42
Figure I.14. Shadowgraph of propellant gas plume. (1: blast wave of first precursor flow; 2: blast wave of propellant flow; 3: Mach disk; 4: barrel shock; 5: turbulent gas-air interface; 6: turbulent vortex ring; 7: increased gas density at downstream propellant gas plume) [17].....	43
Figure I.15. Development of primary, intermediate, and secondary flashes. [17]	45
Figure I.16. Dark-schlieren photograph of muzzle flowfield before projectile ejection (1: blast wave of first precursor flow; 2: turbulent vortex ring) [17].....	50

Figure I.17. Shock wave geometry [25].....	51
Figure I.18. Shock wave reflections [25].....	52
Figure I.19 Gunshot acoustic recording from two microphones emphasizing time difference of arrival of signals (1: shock wave and reflected shockwave; 2: shock wave at near and far microphone; 3: shock wave and muzzle blast; 4: muzzle blast and reflected muzzle blast; 5: reflected muzzle blast at near and far microphones) [24].....	53
Figure I.20. An example model of supersonic weapon fire geometry [25]	54
Figure II.1. Quick reference guide for the pre-detection scheme used.	65
Figure II.2. Quick reference guide for the detection scheme used.	66
Figure II.3. An illustrative example depiction of the pre-detection evaluation process, which assesses pixels that deviate from the statistical background after a period of time.....	69
Figure II.4. An illustrative example of a hostile fire signature depicting the width measured at 10% and 50%.	72
Figure II.5. (Above) A hostile fire signature. (Below) An illustrative example depicting the concept of temporally and intensity normalized signatures (above).	73
Figure II.6. Probability of detection plotted against frame rates for MWIR-1, MWIR-2, and LWIR broadband hostile fire data.....	80
Figure II.7. Detected false alarms for varying frame rates. Low signal noise and clutter, and also reduced samples for background statistics and detection affect the false alarm rates.	81
Figure II.8. Detected false alarms for varying frame rates. Low signal noise and clutter are removed.....	82
Figure II.9. Classification scheme “dynamic rejection” incorporating dynamic false alarm class updating.....	85
Figure II.10. Classification scheme “layered distance” to include false alarms as an additional class by measuring the distance from the respective class.	86
Figure III.1. An example of hostile fire signatures.	90
Figure III.2. Diagram depicting pre-detection memory storage process using an n -window memory system, where $n = 2$, (1: primary and 2: backup). Here, the second detection is greater than the minimum duration threshold. Subsequent detections are overwritten in the backup window.....	93

Figure III.3. A depiction of the pre-detection evaluation process, which assesses the pixels that deviate from the statistical background after a period of time.....	94
Figure III.4. An illustrative example of a temporally and intensity normalized signature.....	96
Figure III.5. Quick reference guide for the pre-detection scheme used.....	98
Figure III.6. Quick reference guide for the detection scheme used.....	99
Figure III.7. Dynamic Distance classification scheme.....	102
Figure III.8. The distance threshold is determined dynamically using the static hostile fire classes and the dynamic false alarm class.....	105
Figure III.9. An example of down-sampling and the choice of offset.....	106
Figure III.10. The probability of detection as a function of frame rate for LWIR, MWIR-1, MWIR-2, and LWIR/MWIR-1.....	107
Figure III.11. False detections as a function of frame rate for LWIR and MWIR-1.....	108
Figure III.12. False detections as a function of frame rate for LWIR and MWIR-1 for stationary data.....	109
Figure III.13. 120 Hz data. Probability of hostile fire declaration (black percentage) and false alarms (red numbers) for various architecture configurations, including dual band detection and classification. Classification is performed using the dynamic distance classification scheme...	111
Figure III.14. 600 Hz data. Probability of hostile fire declaration (black percentage) and false alarms (red numbers) for various architecture configurations including dual band detection and classification. Classification is performed using the dynamic distance classification scheme...	111
Figure III.15. 600 Hz and 120 Hz MWIR-2 data. Probability of hostile fire declaration and false alarms for various architecture configurations for single band detection and classification. Classification is performed using the dynamic distance classification scheme.....	112
Figure IV.1. Detection Software Diagram.....	125
Figure IV.2 Detection decision diagram.....	126
Figure IV.3. Classification decision diagram.....	127
Figure IV.4. Weapon fire detection system processing using GPS for sun detection.....	134
Figure IV.5. Sun Detection System.....	136
Figure IV.6. Sun Detection.....	137

Figure IV.7. Sun Locating	138
Figure IV.8. Sun Tracking	139
Figure IV.9. Sun Orientation	140

I. INTRODUCTION

Objective

This work emphasizes methods to detect impulses using signal processing and system design, for real-world application and implementation. The impulses under consideration in this work relate to the high intensity, transient events associated with weapon fire. In particular, the electro-optical and infrared signatures (EO/IR) of muzzle flashes from weapon fire are studied. Similarly, weapon fire impacts and acoustic signatures can also exhibit high intensity, short duration impulse-like signals. Though weapon fire detection is studied in this work, other applications for impulse detection could be heart pulse, lightning strike, or seismogram monitoring and characterization.

The techniques presented in this work provide practical methods for impulse detection which can be used for weapon fire detection. The algorithm allows for detection of not only high intensity, but also low intensity signals, corresponding to near and far range targets. Techniques which rely only on signature intensity levels are often range limited due to background clutter, and can be potentially fooled. Background clutter becomes increasingly burdensome for on-the-move applications such as detection from a moving vehicle. Due to the increased clutter environment when operating from a vehicle, current detection methods often limit the scope of applicability, such as only detecting particular weapons, and only at limited ranges. The algorithm proposed, addresses not only detection, but also clutter rejection for on-the-move scenarios. Because clutter sources can change from one environment to another, the algorithm is an adaptive clutter rejection methodology. Additionally, the algorithm relies on temporal signature profiles rather than simply signal intensity. The temporal profiles allow for detection at lower intensity levels, or increased range, and are less susceptible to modification. Furthermore, pixel level algorithm performance allows detection localization to within the pixel's individual,

or instantaneous, field of view. The generality of the algorithm enables the capability to detect and locate various weapon types, at various ranges, both stationary, and on-the-move.

An algorithm, however, does not operate in a vacuum, but is designed to work in a system. Due to hardware limitations such as available dynamic range in sensors, saturation can occur, which increases the clutter environment. A system design is also provided to reduce limitations of hardware and algorithms, and exploit strengths of the algorithms. Though the algorithms cannot distinguish one saturated signal from another, a system design which provides unsaturated signals to be processed is explored. Multiple designs are offered, which either provide an unsaturated signal or alternative information which identifies the signal source, in order to combat saturation. By the same token, increased information can allow signals to be more easily discriminated. By designing a system, such as a multi-band or multi-modal system, with increased information, the algorithm has increased performance. In some cases, such as incorporating an acoustic sensor, additional information can be provided beyond detection and clutter mitigation. In the case of acoustic sensors, range estimation can be provided, allowing for the additional hardware to be multi-purposed.

The techniques used in the algorithm allow for real time processing. A layered feature extraction and thresholding approach is used for detection of weapon fire and clutter rejection. The same features measured for the detection process, are used by classifiers to provide signature classification. Thus, in the case of weapon fire detection, the algorithm provides a weapon class as an output. The classification, however, is also used as an additional means of clutter rejection, beyond the initial detection stage. Additionally, the classification scheme is periodically updated to reflect changes in the local background, enabling practical use in multiple and unpredicted

environments. The processing performed has been made to run in real-time, making it a practical and useful approach to weapon fire detection.

The resulting algorithm is shown to enable detection of weapon fire targets, while minimizing false alarms, for real-world, field-able applications. The work presented demonstrates the complexity of detection algorithm and system design for practical applications in complex environments, including signature analysis, algorithm analysis, and system design. The work presented also emphasizes the complex interactions and considerations when designing a practical system, where system design is the intersection of algorithm performance, processing, and design, hardware performance and design, and size, weight, power, and cost (SWAP-C) (see Figure I.1). The trade space is illustrated by example in considering design components of sensor parameters, algorithm complexity, and hardware. Sensor parameters can affect signal-to-noise (detection, range performance), and environmental clutter (false alarms). Algorithm complexity is dependent on detection rate, detection range performance, and false alarm mitigation. The algorithm must extract enough information to make a reliable decision. More sensors or multi-modal sensing can increase the amount of information available. Increased algorithm complexity can increase the amount of information extracted. Algorithm complexity affects real-time processing requirements (processing speed, power, and heat). Practical design is affected by cost, hardware/sensor availability, and size, weight, and power (SWAP). The algorithm design presented in this dissertation takes into account system detection and false alarm mitigation strategies, using available hardware, in real world applications and operational use.

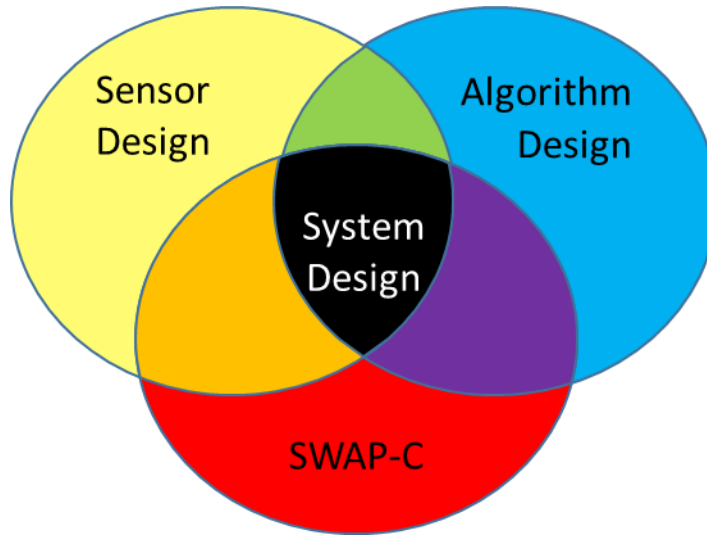


Figure I.1. Diagram depicting system design as the intersection of sensor design, algorithm design, and SWAP-C

Hypothesis

The algorithm and subsequent systems are intended to further improve upon the state of the art by providing practical, functional impulse target detection systems and methodologies which take into account detection performance, false alarm and saturation clutter mitigation, cost, hardware availability, range performance, on-the-move operation, and real-time processing, thus ensuring the transition from lab demonstration to real world application.

Novelty

The work presented here improves upon the state of the art by providing a generalized impulse detection algorithm which can be applied to EO/IR sensors, independent of waveband. When applied to weapon fire detection, the algorithm, and subsequent system designs, allow for on-the-move detection and classification of various weapon fire types, at various ranges, in real-time, with minimized clutter.

The developed algorithm has enabled target detection in multiple spectral bands in the EO/IR. These bands exhibit different clutter profiles, especially when the system is in motion.

The complexity of the algorithm, to include an adaptable clutter rejection and classification scheme, allows for detection of targets in complex environments, while minimizing false alarms. The algorithm was also shown to work at various video frame rates of operation. Traditionally, image-based detection of weapon fire has operated at higher frame rates and the proposed algorithm exceeds the state of the art by allowing low frame rates in a variety of wavebands. The versatility of the algorithm to operate in multiple wave bands and frame rates, allows flexibility in system design, which takes into account real world aspects of cost, hardware availability, and range performance.

The algorithm, which takes into account, not only intensity, but duration and shape, is able to extract more discriminating information, allowing for increased target detection for low intensity, far range targets. Short range, high intensity targets can generate saturated signals, a limitation of sensor hardware to detect signal from both far and short-range targets. Saturated targets can be indistinguishable from saturated clutter. System designs are provided which address saturation by providing means of detecting saturated signals and reducing saturation clutter, through the use of multiple wavebands and multi-modal sensing. The system design improves upon the state of the art by addressing saturation in system designs through the use of EO/IR saturation detectors, acoustic saturation detectors, and position, navigation, and timing (PNT) sensors to limit saturation clutter.

Additional system designs are provided which exploit the strengths of the algorithm performance. Complex acoustic systems already exist for weapon fire detection. When acoustic sensing is combined with the proposed EO/IR algorithm and system, the acoustic sensor can simply be used as a confirmation sensor, reducing restrictions of acoustic sensing complexity. A system is presented which incorporates PNT data, in order to minimize saturation clutter,

ultimately enabling greater detection of near range targets, complementing the algorithm's far target range detection performance. Additional potential functionality is described that can be achieved in addition to target detection and clutter rejection, by the addition of both acoustic sensing and PNT data.

Author Contributions

The following section specifies the contributions of the author contained in this work. The contributions span across supporting background technical literature and techniques, algorithm process and evaluation, and system design. The contributions will be provided for each chapter.

In Chapter 1, the author speaks to the concept of determining a representative background for target to background analysis. The author presents an original custom method for auto determining the target background, based on target dimensions, and the location of the target within the image. This method is intended for practical, low computational complexity applications. Equations are listed which were derived by the author. Additionally, in Chapter I, "Introduction – Multi-Frame Image Analysis," the author provides a discussion of running statistics. Because these statistics require a number of samples before the resulting data is statistically representative of the background, and also because these statistics can "age-out" due to an ever-increasing past history of samples included in the statistics, the author presents an original method to continually refresh the generated statistics, labeled alternating staggered windows.

In Chapter 2, the author presents an original process and assesses techniques to indicate, extract, and detect weapon fire. Additionally, the author provides an original assessment of false alarms and original features for performing classification, with the purpose of classifying targets,

as well as reducing false alarms. Specifically, the weapon fire declaration process is divided into a pre-detection stage, a detection stage, and a classification stage.

The pre-detection process uses techniques such as image pre-processing, statistical background and buffer, and target thresholding to indicate pre-detections. The original process, presented by the author, includes pre-processing to reduce motion effects, pre-detection signal indication, tracking, evaluation of signal after a duration, peak pixel clustering or grouping. Though individual components of the process, such as maximum filtering, sliding statistical window, and maximum pixel intensity are not unique, the combination of techniques used to address reliable weapon fire signal pre-detection with consideration for false alarm reduction, data reduction, and on-the-move application is an original contribution. Similarly, frequency filtering is not a unique concept, however, frequency filtering weapon fire events to reduce false alarms, as well as reduce data storage and processing, is an original contribution in this work.

The detection process consists of signature extraction and feature extraction. All techniques discussed for weapon fire signature extraction are original contributions in this work. Though signature analysis is often discussed, the process for extracting weapon fire signatures for reliably consistent temporal profiles has not been discussed in other works, to the author's knowledge. The extracted signature is then measured to determine if a weapon fire detection occurs. The 50% and 10% width measurement and rise time measurement techniques have been presented in other works for a subset of weapon types. The demonstration of the 50% and 10% width measurement and rise time measurement to an alternate set of weapon types is an original contribution of this work. Other techniques to measure duration were assessed, such calculating the area under the curve and measuring standard deviations with the assumption of a Gaussian distribution. These two additional duration measurement techniques for the application of

weapon fire is an original contribution. Additionally, the fall time measurement as a method for discriminating weapon fire is an original contribution. The technique of sampling and normalizing weapon fire into 10 samples and all subsequent analysis performed on the normalized samples, are original contributions in this work.

The detection process also investigates the detected pixel's surrounding neighborhood of pixels. The analysis techniques of co-variance, correlation, median filtering, summation, and combinations thereof to assess the surrounding pixel neighborhood are original contributions. The weapon fire detection process, to include signature extraction and sequential signature measurement to not only determine weapon fire, but also reduce false alarms and data processing is an original contribution.

For the classification stage, the analysis of measurements for their suitability to be used as features which provide the greatest degree of class separation is an original contribution. Once again, the use of the normalized intensity samples, and specifically, their use as features for classification, is an original contribution. The spectral analysis of measurements and their ability to separate classes in respective spectral bands, is also an original contribution.

The results presented in Chapter 2 are original contributions. The results include a breakdown of sequential detection measurements and their ability to reduce false alarms, per spectral band. The results also include a clutter source analysis, per spectral band. The results also present a framerate analysis per spectral band. The framerate analysis demonstrates detection capability at 120 Hz, and even 100 Hz. This work establishes the ability to detect certain types of weapon fire at 120 Hz in the field of weapon fire detection. The work not only investigates detection, but also false alarms at varying framerates, per spectral band.

In Chapter 2, a series of processes are assessed for the purposes of performing classification, and in some cases, reducing false alarms. These processes are original contributions. Though common classifiers are used, the analysis of their application to the data set, and their sequencing in multi-layer classification schemes is an original contribution. Additionally, two classification schemes are provided for the specific purpose of reducing false alarms. These two classification schemes are original contributions. The distance metric used to evaluate observations with respect to their respective labeled class is an original contribution.

Chapter 3 extends upon the work of Chapter 2. Chapter 3 adds additional methods for data reduction, signature extraction, multi-band detection, cross-band spatial and temporal search windows, and a scheme for classification and false alarm reduction. The n-window system is an original contribution and is used to limit memory usage and reduce the number of pre-detections processed for evaluation, allowing for real time data processing. As in Chapter 2, all signature extraction techniques for weapon fire are original contributions. Though multi-band detection is a common technique, the presented process for performing multi-band detection, to include temporal and spatial cross-band search windows to allow for reduced system design tolerances is an original contribution. The classification scheme, presented in Chapter 3, to include class cluster distance, feature distance, and distance threshold determination, are all original contributions.

The results contained in Chapter 3 address frame rate analysis, classifier performance, and system performance, per spectral band, as well as for dual band application. The results also provide an analysis of false alarms per spectral band and dual band, based on weather conditions and also scene motion. The results provide a confusion matrix which details classification performance given the data set and selected features. These are all original contributions.

Chapter 4 generalizes the methods provided in Chapters 1 and 2. Chapter 4 provides flowchart representation of the systems and algorithmic process contained therein. The method described in this chapter has been published by the United States Patent and Trademark Office under patent number US 10,389,928 B2. The US patent is an original contribution, by definition.

Chapter 4 presents additional system designs and their associated methodologies intended to reduce false alarms and address saturation. In one design, a tri-band detection system is presented. Multi-band detection is a common approach. The use of a tri-band system, however, to not only reduce false alarms, but also to reduce saturation clutter, is an original contribution. Another presented design extends upon this tri-band system, to include a dedicated saturation sensor. The addition of a specific saturation sensor is an original contribution. A third design and original contribution reuses the idea of a specific saturation sensor, where in this case, the saturation sensor is an acoustic sensor. Though the addition of an acoustic sensor to allow multi-modal sensing is not a novel technique, the use of an acoustic sensor to detect events which saturate optical imagers is a novel and original technique. A fourth design incorporates a sun localization method. The method for determining the sun's location is a common technique. The idea of using the sun's known location within the image to remove saturation caused by direct view of the sun is an original contribution. The systems described in this chapter have been published by the United States Patent and Trademark Office under patent number US 10,209,343 B1.

Chapter 4 expands upon previously discussed methods to include an acoustic sensor for multi-modal detection. Multi-modal sensing for weapon fire detection, as stated previously, is a known technique. The method described in Chapter 4 emphasizes clutter mitigation by using the acoustic sensor as a confirmation sensor to the EO/IR detection sensor. This is an original

contribution in that the well-known multi-modal approach for weapon fire detection uses an EO/IR sensor and a confirmation sensor to an acoustic detection sensor. The system and method described in this chapter has been filed for application at the United States Patent and Trademark Office on August 9, 2018 under application number 15/998,345.

Background

Detection Theory

Detection

In [1], a history and basic elements of detection theory are described, and the reader is directed to this work for a more detailed detection theory description. The basic concept of detection is the discrimination between two stimuli, where one stimulus contains the null stimulus. An example of this concept is the discrimination of a target from a background. The discrimination performance is often measured in terms of sensitivity, such that high sensitivity measure indicates a better ability to discriminate the target from the background, and a lower sensitivity indicates a poorer ability to discriminate the target from the background. In describing discrimination and sensitivity, four measures or probabilities are often used, the hit rate (p_H), the false alarm rate (p_F), the miss rate (p_M), and the correct rejection rate (p_R). The hit rate is the ratio of the number of correctly discriminated targets, given the total number of targets presented. The miss rate is the ratio of the number of targets not discriminated, given the total number of targets presented, and can be represented as the complement of the hit rate, $1 - p_H$. Similarly, the correct rejection rate is ratio of correctly discriminated background presentations to total number of background presentations. The false alarm rate is ratio of background targets, incorrectly discriminated as targets, to the total number of background presentations. The correct rejection ratio, represented as $1 - p_F$, is the complement of the false alarm rate. Because of the

complementary nature, the discrimination task is often analyzed in terms of hit rate and false alarm rate. Alternately, the miss rate and false alarm rate are referred to as type 1 error and type 2 error, respectively. The four probabilities, p_H , p_M , p_F , p_R , can be described as

$$\begin{aligned}
 P[\text{Hit}] &= p_H = \frac{\text{Hit}}{\text{Hit} + \text{Miss}} \\
 P[\text{Miss}] &= p_M = \frac{\text{Miss}}{\text{Hit} + \text{Miss}} = 1 - p_H \\
 P[\text{False Alarm}] &= p_F = \frac{\text{False Alarm}}{\text{False Alarm} + \text{Correct Rejection}} \\
 P[\text{Correct Rejection}] &= p_R = \frac{\text{Correct Rejection}}{\text{False Alarm} + \text{Correct Rejection}} \\
 &= 1 - p_F
 \end{aligned} \tag{1.1}$$

It is desirable for a detection system to be perfectly sensitive to the presence of a target, such that the system always detects targets, and has zero false alarms. In many practical applications, the probability of hitting (p_H) or detecting a target is greater than the probability of false alarm (p_F), but p_F is not zero. It is important, however, to describe the performance of a systems in terms of both p_H and p_F . For example, a system which always says that a target is present, will have a perfect hit rate of 1, but will also have a false alarm rate of 1, indicating poor performance, even though all of the targets were labeled as targets. Acceptable performance of the system with respect to p_H and p_F in practical applications is dependent on the reward or risk associated with detecting or missing a target, and the reward or risk associated with correctly rejecting a target or generating a false alarm. The sensitivity measure, probability of correct responses ($p(c)$), takes into account the hit rate and the correct rejection rate, such that $p(c) = \frac{1}{2}[p_H + (1 - p_F)]$. Alternately, a sensitivity measure of $p_H - p_F$ could also be used. Note that there is an assumption of equal presentation of targets and backgrounds for these sensitivity measures [1]. Both measures, however, provide an indication of the discriminability of the target

from the background. Often, the performance of a system is presented in terms of p_H , p_F , and sensitivity using receiver operating characteristic (ROC) curves on a plot [1].

In detection theory, a sensitivity measure known as d' (“dee-prime”) converts the ratio or probability values of p_H and p_F to a distance in terms of standard deviation units using the inverse of the normal distribution function or z transformation. In a normal distribution, a given distance or value, along the x-axis, corresponds to an associated probability along the y-axis. Conversion of p_H and p_F into distances is preferable when describing the distribution of stimuli or classes, given a parameter or feature, and their ability to be separated or discriminated. A distance measure can be used to describe the separation between two classes, and help to determine a decision threshold criterion for choosing one stimuli versus another. Assuming normal distributions for both stimuli, the distance between the two distributions, along a decision axis, is measured in standard deviations. When the distributions have different standard deviations, the distance can be scaled for its relative importance, or distance from the mean of each class. This is important when trying to determine a decision threshold, for instance, which is equally distant from the two classes (i.e., zero bias). When the standard deviation is the same for both distributions, an equally distant threshold criterion is the midway point between the means of the two distributions. In [1], multiple methods for scaling the decision axis to account for differences in standard deviations are presented. Additionally, an example method is presented here.

An indication of discrimination accuracy can be defined by the probability of classification, given by

$$p_{class} = \frac{Hit + Correct\ Rejection}{Total\ Number\ of\ Stimuli\ Presented} \quad (1.2)$$

In (1.2), the determination of the probability of correct classification is dependent on the base presentation rates. Thus, if more background scenes are presented, a higher classification rate might be achieved by biasing the decision threshold to choose class “background” a greater percentage of times. If we assume that the distributions of background and target can be described as Normal distribution, as depicted in Figure I.2, a threshold criterion can be determined independent of base presentation rate.

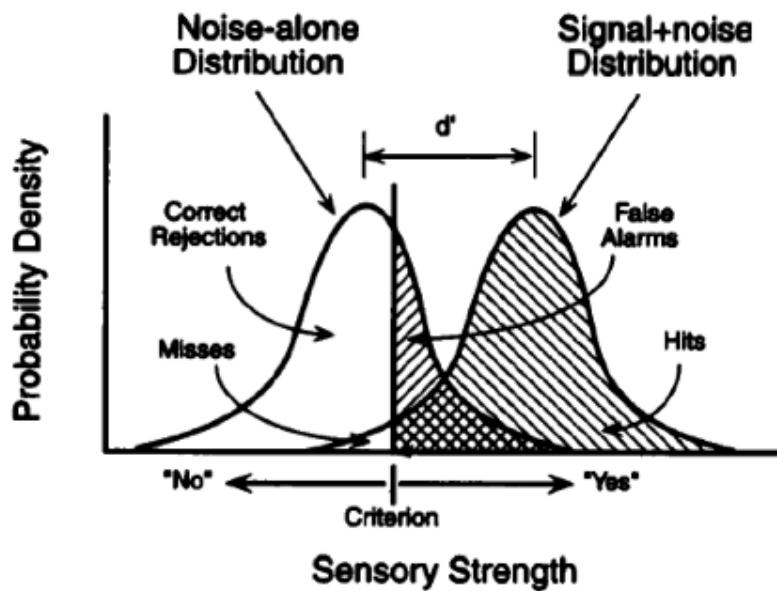


Figure I.2 Signal detection theory definitions and assumptions. [1] N. A. MacMillan and D. C. Creelman, *Detection Theory" A User's Guide*, Mahwah: Lawrence Erlbaum Associates, 2005.

In general, if the two distributions have identical standard deviations, but differing means, an optimal zero-bias threshold would be the midpoint between the two means. When the standard deviations of the two distributions are different, the optimal threshold can still be achieved when the threshold lies equidistant between the two means. The measure of distance, however, is relative to the shape of the distribution, as depicted in Figure I.3. The distance of the threshold, τ , from the distribution S , can then be measured between the mean of S , μ_S , in units

of standard deviations, σ_S , corresponding to distribution S. Given a distribution of target signal, T, and a distribution of background noise, B, the distance can be measured in units of standard deviation, or, alternatively corresponding z-scores generated by the z transformation

$$\begin{aligned} z(p_H) &= \frac{\mu_T - \tau}{\sigma_T} \\ z(p_F) &= \frac{\mu_B - \tau}{\sigma_B} \end{aligned} \tag{1.3}$$

By setting the distance values in (1.3) to be equal, an optimal criterion threshold can be found by solving for τ , and is given in (1.4).

$$\tau = \frac{1}{\sigma_T + \sigma_B} (\sigma_B \mu_T + \sigma_T \mu_B) \tag{1.4}$$

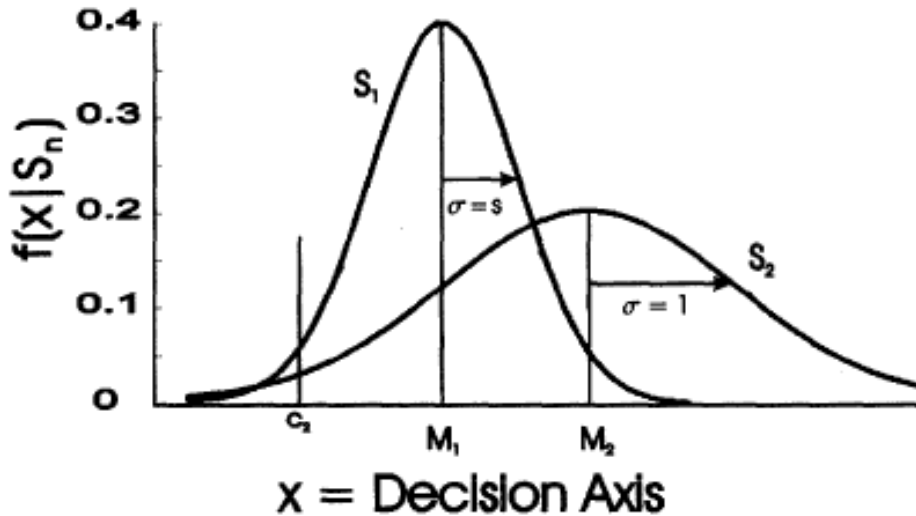


Figure I.3. Distributions with unequal standard deviations. [1] N. A. MacMillan and D. C. Creelman, "Detection Theory" A User's Guide, Mahwah: Lawrence Erlbaum Associates, 2005.

Classification

The process of sorting multiple stimuli into a smaller number of classes, is the task of classification [1]. The classification process typically consists of data collection, feature choice,

model choice, training, and evaluation [2]. The choice of features is dependent on prior knowledge, and in general, the features should be easy to extract, invariant to irrelevant transformations, insensitive to noise, and useful for discriminating patterns in different categories or classes [2]. In evaluating the classification system for practical use, it is important that the system is able to classify the training data correctly. The system should also avoid over-fitting, in which case, the classifier is not generalized enough to correctly classify new unseen samples. Variety in data collection, and thus training data, can help to avoid over-fitting. Additionally, the computational complexity should be taken into account for the application. In many cases, the classifier is trained in the lab, while the decision process occurs in the field. Thus, the complexity of learning can be high, while the decision complexity is of concern. In other cases, learning and decision making can both occur in the field, and the computational complexity of both should be taken into account for a fielded system.

A fundamental statistical approach to classification is Bayesian decision theory [2]. Given an observation, features are measured from the observation, and the observation is subsequently classified into one of two or more classes. Let $\{\omega_1, \dots, \omega_c\}$ denote a set of c classes, where ω_j is the j^{th} class within the set. The decision to classify the observation into one of the classes, for example, the j^{th} class, is based on prior knowledge about the expectation of the frequency of occurrence the class (ω_j) , denoted by $P(\omega_j)$, and the likelihood, that given a class (ω_j) , a particular measured feature (x) will be observed, denoted by $p(x|\omega_j)$. This, in turn, is used to determine the posterior probability, denoted by $P(\omega_j|x)$, that given an observed measured feature, a particular class was observed. The decision to classify an observation as a particular class is governed by the posterior probability and the risk or cost associated with an incorrect classification. The probabilities, however, are generally not known,

as they represent the entire population of data. In a typical scenario, only a sample of the total possible population data is available, and the prior probability and likelihood are inferred by the sample data set. Thus, sample data or training data, and its representation of the population data set, affects the ability to accurately classify. The underlying probability density function is often assumed, such that only the parameters are estimated from the training sample data. For example, the Normal distribution is completely described by the parameters of mean and variance, and the parameters are estimated from the data.

Other methods exist for determining decision boundaries for classifying observations into classes. Linear discriminant functions are weighted linear combinations of measured features, where each class has an associated linear discriminant function. The decision to choose a particular class is based on the discriminant function providing the highest value. The value of the discriminant function is generated by adding the weighted contributions of each feature. In this method, the weights are determined from the training data, such that the classification error is minimized. For further reading on methods for determining weights, the reader is directed to [2]. Quadratic discriminant functions are linear discriminant functions with the addition of weighted feature pairs. Thus, a measure of how the features co-vary together, for a particular class, is taken into account. Support vector machines (SVM) are a form of linear discriminant function. In SVM, the features are transformed to a higher dimensional space, and the transformed features are added together as a weighted linear combination. The feature space is transformed to a higher dimension for the purpose of separating classes. At lower dimensions, the classes may exhibit complex multi-point decision boundary. By increasing the dimensions of the feature space, a single hyperplane decision boundary is found, equally distant from either class, where the class boundary is indicated by the support vectors. The separation is increased,

not only for accurate classification of training data, but also for robustness of correctly classifying new unknown observation. It should be noted that SVM provides classification for the binary case of two classes, where linear discriminant function analysis (LDA) provides classification for two or more classes.

Though general principals of detection and classification methods were discussed, practical methods must be used for determining the presence of stimuli and extracting features for analysis. For images, this entails determining significance in observed pixel intensity differences and subsequently extracting features which describe and distinguish those observed and detected events. The following provides a discussion on methods used to determine significance or targets from background imagery.

Contrast analysis

Contrast is the difference in luminous values (lightness or brightness) of two objects or stimuli that allows the two objects to be distinguishable [3]. The nature of the human visual system is such that the perceived contrast is a function of not only luminance difference but also spatial frequency. Thus, contrast is perceived differently for different spatial frequencies. This property is evident in Figure I.4. Here, contrast increases monotonically from the top of the figure to the bottom of the figure, along the vertical axis, while frequency increases from left to right, along the horizontal axis. The upper portion of the figure shows that perceived contrast is lower for higher spatial frequencies and also for very low frequencies.

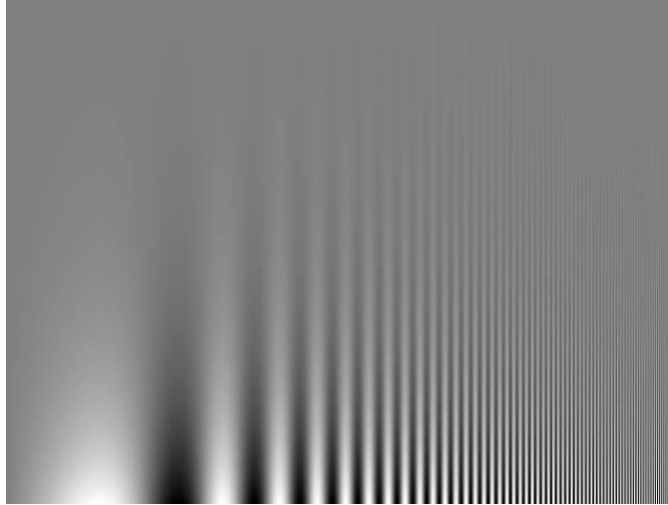


Figure I.4. Contrast vs spatial frequency. Contrast is varied vertically. Spatial frequency is varied horizontally.

Image contrast can be defined as

$$C_1 = \frac{\textit{Luminance Difference}}{\textit{Average Luminance}} \quad (1.5)$$

where the luminance difference is measured between features in the image and the average luminance is the set which includes the features being measured. For a digital image, the luminance difference can be interpreted as the intensity difference between the maximum and minimum values in the image, and the average luminance is the average pixel intensity of the image. In the same tone, the contrast between a feature and the background would be defined by the luminance difference between the feature and the background normalized by the average luminance of the background and the feature. If the feature is small in comparison to the background, the average luminance can be approximated as the background luminance. This approximation of the average luminance in (1.5) leads to the Weber contrast metric which is defined as

$$C_2 = \frac{L_{Tgt} - L_{Bgd}}{L_{Bgd}} \quad (1.6)$$

where L_{Tgt} and L_{Bgd} are the luminance values describing the feature (target) and the background, respectively [4]. The Weber contrast implies that small changes between a target and background are more noticeable when the background luminance is small. This metric for evaluating contrast holds when there is a single target of uniform luminance against a background of uniform luminance [4]. Another interpretation is that the viewer is assumed to be adapted to the background and is only viewing changes with respect to the background [5]. The metric may not hold if the background is not uniform, such that it consists of multiple targets or excessive noise. Additionally, the Weber contrast values lie in the range of -1.0 to $+\infty$, where metrics ranges are typically bounded.

An alternate metric for evaluating contrast is the Michelson contrast which is defined as

$$C_3 = \frac{L_{Max} - L_{Min}}{L_{Max} + L_{Min}} \quad (1.7)$$

where L_{Max} and L_{Min} are the maximum and minimum luminance value of an image. This metric assumes that the viewer is presented with equal amounts of foreground stimuli and background. Here, the viewer is adapted to both foreground and background and is tasked with denoting changes. This can be the case if there are multiple targets or periodic patterns such as a sinusoidal grating and thus the background is not uniform [4]. Because of the normalizing factor in (1.7), the range of values for the Michelson contrast is bounded between 0 and +1.0.

The Michelson contrast, however, suffers in that single extreme luminance values in an image can dramatically affect the associated contrast. Two images could appear different but have the same contrast values because only the minimum and maximum luminance values are

used [6]. The root-mean-square (RMS) contrast can be employed to reduce the effect of outliers and is defined as

$$RMS = \left[\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{1/2} \quad (1.8)$$

where for a digital image with pixel values normalized to [0,1], n is the number of pixels, x_i is the intensity of the i^{th} pixel, and \bar{x} is the mean intensity defined by

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1.9)$$

The RMS contrast is essentially the standard deviation of the pixel values in the image. Here, all pixel values are accounted for, and not just the minimum and maximum values. The RMS contrast, in itself, does not account for the spatial frequency content or spatial distribution of contrast in the image [4]. The image can further be described by using (1.8) to measure contrast at specific spatial frequencies to be used as a measure of spatial resolution [6]. Metrics which inherently address spatial frequency can also be considered and the reader is directed to [4].

The metrics of contrast defined in (1.5-1.9) have often been used or modified to fit specific applications. It is common to consider the contrast of specific targets with that of a surrounding background. In this case, the contrast of the entire image may not be considered, but merely the contrast of a region of the image. Or, it may be the case that the image was produced with an imager having a narrow field of view, such that all targets, or sets of targets, at a particular range are visible. In considering contrast of thermal IR imagers, such as mid-wave infrared (MWIR) and long wave infrared (LWIR) imagers, the Michelson contrast is modified to

reflect the path radiance due to particles in the atmosphere contributing to the overall background radiation viewed by the detector and thus lowering the contrast [7]. This contrast is given by

$$C_4 = \frac{L_{Tgt} - L_{Bgd}}{L_{Tgt} + L_{Bgd} + 2L_{Path}} \quad (1.10)$$

where L_{Tgt} and L_{Bgd} are the transmitted target and background radiance and L_{Path} is the atmospheric path radiance as seen by the sensor. This metric assumes that L_{Tgt} and L_{Bgd} , which do not contain contributions from path radiance, are known. The contrast defined by (1.7) was further extended to reflect that the path radiance may not be static with time, such as when there is a turbulent dust cloud in the atmosphere [8]. In this LWIR analysis, the additive portion of the degradation caused by the dust is modeled into the contrast function as

$$C_5 = \frac{L_{Tgt} - L_{Bgd}}{L_{Tgt} + L_{Bgd} + 2L_{Path}(r, t)} \quad (1.11)$$

where (r,t) denotes the dependency on range distance to the target and time. In (1.10), the path radiance could be scaled out to enhance the contrast of targets at a given range using the equation

$$C_6 = \frac{L_{Tgt} - L_{Bgd}}{L_{Tgt} + L_{Bgd} - \beta} \quad (1.12)$$

Here, the path radiance is included in the L_{Tgt} and L_{Bgd} to reflect added path radiance and transmission loss. The term, β , is used to counteract the reduction in contrast due to path radiance [7]. If the path radiance is a function of time, as proposed in (1.11), a static β term may not be useful in consistently removing the effect of the path radiance.

Background analysis

In all cases presented here for determining contrast, the target intensity and background intensity are needed. Establishing these values can be accomplished in different fashions. Targets can be ascertained automatically or manually. Targets chosen automatically typically require

knowledge of the background. If the background is yet to be determined, then it is likely the target will either need to be defined manually or determined through an alternate feature analysis. An example of an alternate feature analysis is detection of all pixels which exceed a given threshold intensity value. An image mask with the target region of interest (ROI) can be used to identify the target in the image. Given the target is detected, the background corresponding to the detected target can be found. The target is distinguished from the background due to its contrast with the surrounding background. Thus, the background is defined as the area surrounding the target, which is of equal size in pixels as the target [9]. The term “area surrounding the target” is not clearly defined, however.

One method of establishing the surrounding background is to simply expand the target area in all directions until the background area is equivalent to the target area. When the target is assumed to be rectangular, this amounts to defining a larger rectangle encompassing the target. When the target is of irregular shape, the process can still be performed, but the resulting background will not be rectangular, but instead will have the irregular shape of the target, scaled. This scenario is depicted in Figure I.5 and Figure I.6. Thus, a background perimeter is formed around the target, and the thickness of this perimeter satisfies the restraint for the ratio of target area to background area being equal to one. This ratio follows the Michelson contrast concept of providing equal stimuli with equal amounts of background. Alternative contrast schemes may call for ratios other than one.

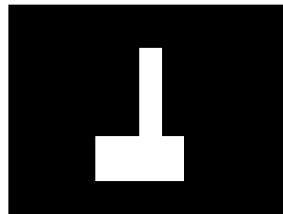


Figure I.5. Irregular shaped target mask

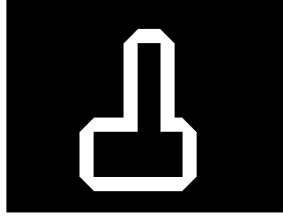


Figure I.6. Background mask corresponding to Figure I.5 using a process of perimeter thickening

The process of thickening the target perimeter to form the background has both pros and cons. An approximate constant thickness surrounding the target implies that local target boundaries are compared against approximately equivalent amounts of background, regardless of the location of the boundary along the target. A “con,” or more precisely, a consideration of the perimeter thickening method is that multiple loops of neighboring pixel evaluations may need to be performed before a solution is determined. Additionally, the background area is defined by an image mask or a list of background pixels. For remote real time applications, processing power and time needed to assess the background ROI, as well as memory storage, should be evaluated.

An alternative approach is to assume the target can be approximated as a rectangular or elliptical shape. In this scenario, the background can then be approximated as a bounding rectangular box. Using the bounding box approach to determining background produces a result that can be solved for directly, in most cases, and the result can be defined by two sets of coordinates. This approach is both faster and simpler than the perimeter thickening method, but relies on assumptions about the shape of the target. Additional restrictions can be placed on the bounding box background, such that the *height/width* background ratio is equivalent to the *height/width* target ratio. Solving the size and dimension restriction equations produces a value for the dimensions of the background. For a target with a height, h_{tgt} , and width, w_{tgt} , as shown in Figure I.7, the background height, h_{bgd} , and width, w_{bgd} , are defined as

$$h_{bgd} = \sqrt{2 * Pix_{tgt} * (h_{tgt}/w_{tgt})} \quad (1.13)$$

$$w_{bgd} = \sqrt{2 * Pix_{tgt} * (w_{tgt}/h_{tgt})} \quad (1.14)$$

Here, Pix_{tgt} is the total number of pixels in the target. The target and background, along with corresponding dimensions, are shown in Figure I.7.

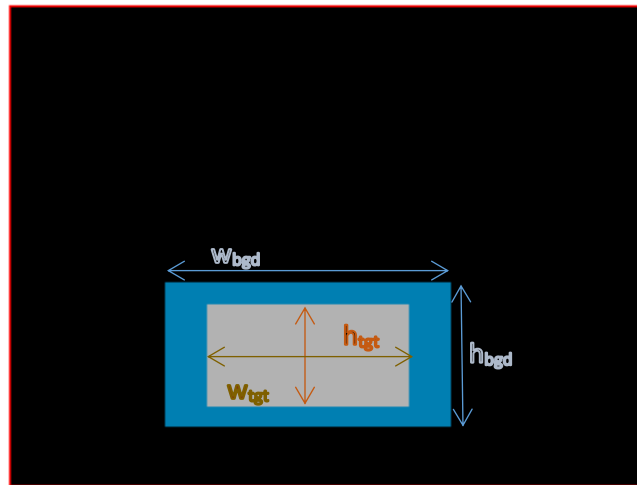


Figure I.7. Target (grey) with bounding box background (blue) of similar proportion and equal size

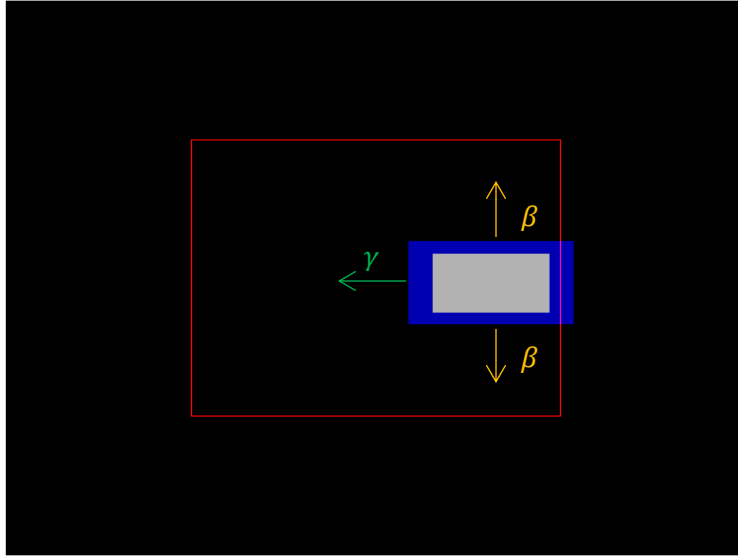


Figure I.8. Target (grey) with bounding box background (blue), partially outside of the image border (red)

The bounding box method, an original contribution of this dissertation, can directly produce dimensions for the background ROI from target pixel parameters, without using an iterative approach. Even when the assumption of the rectangular target holds, the auto determined background may fall out of bounds of the actual image dimensions, however. This is the case when the target lies near the image border, shown in Figure I.8. If satisfying the restriction of the background area being equal to the target area is a priority, then the out of bounds background area can be redistributed within the image boundaries. Thus, if a portion of the initial bounding box is found to lie outside the image border, $Excess_{bgd}$, the boundaries of the generated background that lie within the image borders can be extended to accommodate the size restriction. Suppose we let $InBound = [left, right, top, bottom]$ denote the sides of the background bounding box that lie within the image and thus are eligible for extension. Here the directions (left, right, top, bottom) are equal to 1 when those sides are in bounds and zero elsewhere. For example, Figure I.8 would have $InBound = [1, 0, 1, 1]$. We can then define

terms $\alpha_1 = InBound(1) + InBound(2)$ and $\alpha_2 = InBound(3) + InBound(4)$. If we choose restriction of the redistributed area to maintain the same proportions as the target, we can determine the distance to extend each valid side, denoted by γ and β . Here γ is the amount to extend horizontal valid sides and β is the amount to extend vertical valid sides. The values of γ and β can be found using the positive solution of

$$\gamma, \beta = \frac{-K_{2i} \pm \sqrt{K_{2i}^2 + 4K_{1i}Excess_{bgd}}}{2K_{1i}} \quad (1.15)$$

Where $i = 1$ for γ , and $i = 2$ for β and

$$\begin{aligned} K_{11} &= \alpha_1^2 (h_{tgt}/w_{tgt}) \\ K_{12} &= \alpha_1 h_{bgd} + \alpha_1 w_{bgd} (h_{tgt}/w_{tgt}) \\ K_{21} &= \alpha_2^2 (w_{tgt}/h_{tgt}) \\ K_{22} &= \alpha_2 w_{bgd} + \alpha_2 h_{bgd} (w_{tgt}/h_{tgt}) \end{aligned} \quad (1.16)$$

Here w_{bgd} and h_{bgd} are the horizontal width and vertical height dimensions of the bounding box background which lay within the image boundaries. Using this bounding box boundary adjustment process could result in extended borders lying outside of the image borders. The process can then be repeated until all sides are within the image boundary. In a worst-case scenario, the adjustment process will be repeated three times.

The process of adjusting the bounding box finds a balance between processing and functionality. The perimeter thickening provides increased functionality at the cost of increased processing. The bounding box has minimal processing, but also limited functionality. The adjusted bounding box adds functionality through repeated adjustments, but does not require

element-wise (pixel) evaluations, repeatedly across the image. The adjusted bounding box still does not consider target rotation. Elongated rectangular targets at an angle, may not be correctly evaluated by a bounding box which only considers the target's vertical and horizontal extent. Adjusting the initial bounding box method is trivial. Once again, if the target is approximated by a rectangular or elliptical shape, the major and minor axis dimensions, as well rotation can be found. The major and minor axis dimensions can be used in (1.15) and (1.16) replacing h_{tgt} and w_{tgt} . The resulting bounding box can then be rotated to fit the target. The bounding box is then described by two sets of coordinates and an angle of rotation. Performing the adjusted bounding box procedure using rotation, however, begins to increase processing to the extent that the perimeter thickening may be just as effective. A benefit of the rotated adjusted bounding box would still be the minimal amount of data necessary to describe the shape of the background. It may be useful to check whether a target's dimensions and orientation are such that the target box, defined by the target's horizontal and vertical extent, does not accurately describe the real target. This can be an indication of whether a rotated adjusted bounding box is necessary to describe a target. For a rectangular or elliptical target, we can let the major and minor axis dimensions be defined as Γ and γ and rotation as θ . The un-rotated target bounding box, defined by the target's horizontal and vertical extent, will have an approximate excess area of pixels which does not cover the rotated target, defined by

$$Excess\ Area = (\Gamma^2 + \gamma^2)\cos\theta\sin\theta \quad (1.17)$$

Target Background Distance

The measure of contrast provides a relationship between target and background. This relationship defines the difference which distinguishes a given target from a given background. Contrast measures such as the RMS contrast (1.8), can be thought of as measuring the distance of

a target from its surrounding background. The difference or distance of each target pixel intensity is measured against the average of the background pixel intensities. The average of these distance measures produces the RMS contrast. This technique, however, may not be applicable when the background has large amounts of variation. A pixel value with a large distance from the background mean may not be significant when the background varies a large distance from the mean itself. This point is illustrated in Figure I.9 and Figure I.10.

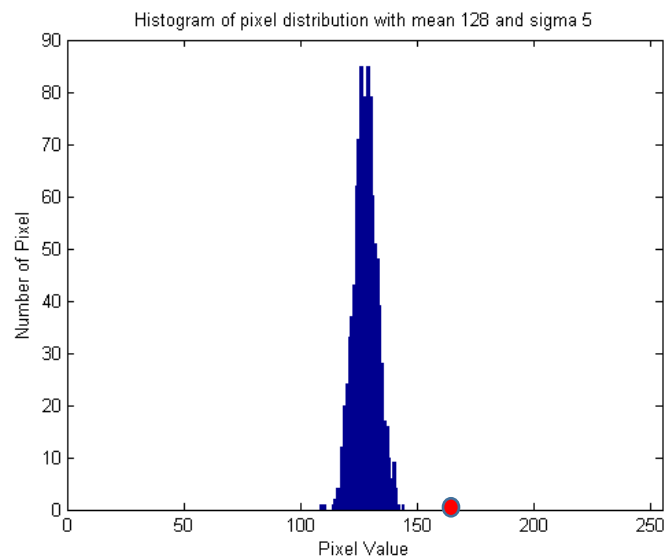


Figure I.9. Histogram of background with $\mu = 128$ and $\sigma = 5$. Target pixel (red) with distance of 30 from the background μ is distinct from the background.

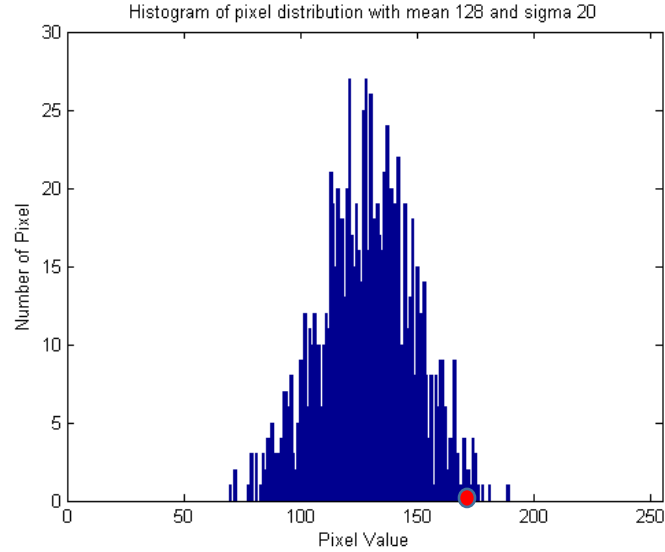


Figure I.10. Histogram of background with $\mu= 128$ and $\sigma = 20$. Target pixel (red) with distance of 30 from the background μ is not distinct from the background.

To accommodate for the variation in the background, the distance measure can be normalized by the variance of the background. This provides for a scale invariant measure of distance, as shown in the following equation

$$\text{Normalised Euclidean Distance} = \left[\sum_{i=1}^n \frac{(y_i - \bar{x})^2}{\sigma_x^2} \right]^{1/2} \quad (1.18)$$

Where y_i is the i^{th} target pixel of n pixels, \bar{x} is the background mean, and σ_x^2 is the variance of the background. This procedure, which calculates the normalized Euclidean distance (NED), can be interpreted as normalizing the data to zero mean and unit standard deviation, and determining the distance from the standard normal of the background.

The equation in (1.18) is useful when the data has only one dimension or feature. This is true when pixel intensity values are used to determine the target distance or contrast from the background. Multiple features may be used however in determining contrast, as in the case of color imagery. For images defined in the RGB color space, each pixel has three features and is

defined by the corresponding intensities in the red, green, blue channels. Thus, each pixel can be described as $x_i = \{x_{iR}, x_{iG}, x_{iB}\}$. Similarly, the mean and variance can be denoted as $\bar{x} = \{\bar{x}_R, \bar{x}_G, \bar{x}_B\}$ and $\sigma_x^2 = \{\sigma_{xR}^2, \sigma_{xG}^2, \sigma_{xB}^2\}$, respectively. And for three channel color images, the NED in (1.18) can be described by

D – Feature Normalised Euclidean Distance

$$= \left[\sum_{i=1}^n \sum_{k=1}^d \frac{(y_{ik} - \bar{x}_k)^2}{\sigma_{xk}^2} \right]^{1/2} \quad (1.19)$$

where k corresponds to the k^{th} of d total features of each observation or pixel.

The d-feature normalized Euclidean distance (dNED) assumes that the features are independent. In a natural image, however, the color of neighboring pixels of an object is likely to be correlated. If the channels are correlated, they vary together. An example of the phenomenon can be seen in Figure I.11 and Figure I.12. In Figure I.11, the two features are independent. This in turn can be interpreted as their covariance is zero and the variance lies in the direction of the feature axes. Therefore, measuring distance along the axis is still appropriate. In Figure I.12, the two features are dependent and vary together. Variance along the feature axes does not completely describe the nature of the distribution. The variance along the diagonals between each pair of feature space axes must also be taken into consideration. By evaluating the covariance of the distribution, a more complete estimate of the shape of the distribution can be used to measure distances from said distribution.

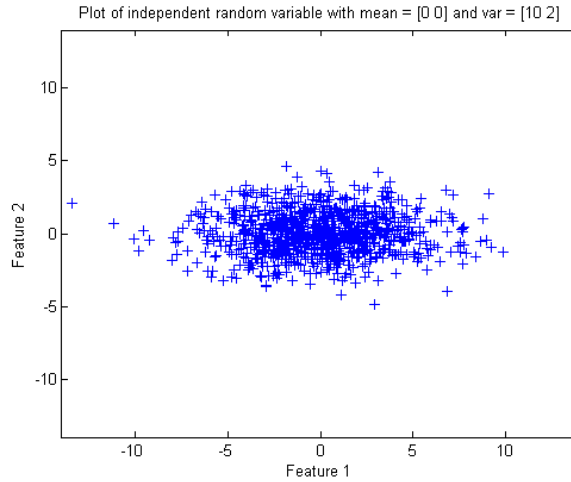


Figure I.11. Distribution of sample observations with two independent features

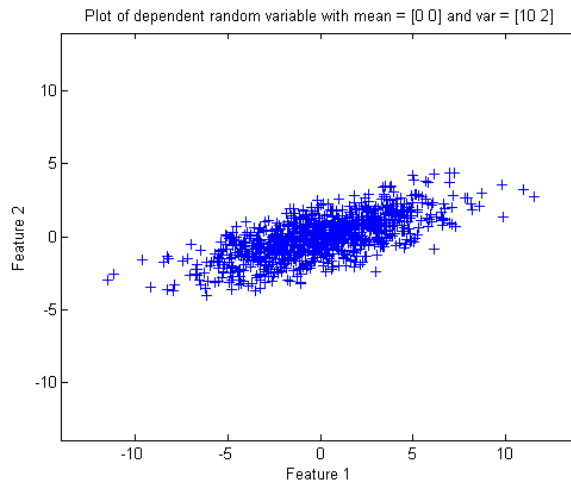


Figure I.12. Distribution of sample observations with two dependent features

Estimating the shape of the distribution completely and accurately is necessary when the multiple dependent features are considered. This, in essence, is the principle behind the distance measure, Mahalanobis Distance, denoted by

$$\text{Mahalanobis Distance } D(y) = [(y - \bar{x})S^{-1}(y - \bar{x})^T]^{1/2} \quad (1.20)$$

where y is the three-dimensional (feature) target data, and \bar{x} is the three-dimensional feature mean of the background (as defined previously), and S is the covariance matrix of the background features. Each entry in the covariance matrix can be computed using

$$\text{Cov}(k_\alpha, k_\beta) = \frac{1}{n-1} \sum_{i=1}^n (k_{\alpha i} - \bar{x}_\alpha) \cdot (k_{\beta i} - \bar{x}_\beta) \quad (1.21)$$

where k_α and k_β are features and $k_{\alpha i}$ and $k_{\beta i}$ are the i^{th} values of n observations of the feature. And \bar{x}_α and \bar{x}_β are the feature means. For the Mahalanobis distance, when y represents a single, three feature, background population mean removed pixel given by $y_i = \{y_{iR}, y_{iG}, y_{iB}\}$, a 1 x 3 matrix, and S^{-1} is a 3 x 3 background feature inverse-covariance matrix, the resulting distance is a single value.

$$\begin{aligned} & \text{Mahalanobis Distance Squared } D(y_i) \\ &= [y_{iR} \quad y_{iG} \quad y_{iB}] \begin{bmatrix} \sigma_{xR}^2 & \text{cov}(R, G) & \text{cov}(R, B) \\ \text{cov}(G, R) & \sigma_{xG}^2 & \text{cov}(G, B) \\ \text{cov}(B, R) & \text{cov}(B, G) & \sigma_{xB}^2 \end{bmatrix} \begin{bmatrix} y_{iR} \\ y_{iG} \\ y_{iB} \end{bmatrix} \end{aligned} \quad (1.22)$$

$$\begin{aligned} & \text{Mahalanobis Distance Squared } D(y_i) \\ &= (y_{iR}^2 \sigma_{xR}^2 + y_{iG} y_{iR} \text{cov}(G, R) + y_{iB} y_{iR} \text{cov}(B, R)) \\ &+ (y_{iR} y_{iG} \text{cov}(R, G) + y_{iG}^2 \sigma_{xG}^2 + y_{iB} y_{iG} \text{cov}(B, G)) \\ &+ (y_{iR} y_{iB} \text{cov}(R, B) + y_{iG} y_{iB} \text{cov}(G, B) + y_{iB}^2 \sigma_{xB}^2) \end{aligned} \quad (1.23)$$

The resulting Mahalanobis distance relates the target pixel intensity to the background. This measure takes into account not only the average background feature values, but also their (co-) variation. Physically, for images, this corresponds to the color pattern of the background or the relationship between pixels in the background. If, however, y represents the $n \times 3$ matrix $y = \{y_1, \dots, y_n\}^T$ (see (1.24)), where there are n three feature target pixels or observations, the resulting Mahalanobis distance is represented by an $n \times n$ matrix.

Mahalanobis Distance Squared $D(y)$

$$= \begin{bmatrix} y_{1R} & y_{1G} & y_{1B} \\ \vdots & \vdots & \vdots \\ y_{iR} & y_{iG} & y_{iB} \\ \vdots & \vdots & \vdots \\ y_{nR} & y_{nG} & y_{nB} \end{bmatrix} \begin{bmatrix} \sigma_{xR}^2 & cov(R, G) & cov(R, B) \\ cov(G, R) & \sigma_{xG}^2 & cov(G, B) \\ cov(B, R) & cov(B, G) & \sigma_{xB}^2 \end{bmatrix} \begin{bmatrix} y_{1R} & y_{iR} & y_{nR} \\ y_{1G} & \dots & y_{iG} & \dots & y_{nG} \\ y_{1B} & y_{iB} & y_{nB} \end{bmatrix} \quad (1.24)$$

$$\text{Mahalanobis Distance Squared } D(y) = \begin{bmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{n1} & \dots & d_{nn} \end{bmatrix} \quad (1.25)$$

In this $n \times n$ distance matrix, shown in (1.25), the diagonal produces the squared Mahalanobis distance $D(y_i) = d_{ii}$, for $i = 1 \dots n$. The off-diagonal values are a measure of the target pixel pairwise relationships normalized by the background covariance matrix. Thus, an increased amount of information can be measured when using multiple observations from the target as it takes into account the physical variation or color pattern of the target with relation to the color pattern of the background. In relationship to the equation presented above in (1.23), each term $y_{i\alpha}y_{i\beta}cov(\alpha, \beta)$ is replaced by $y_{i\alpha}y_{j\beta}cov(\alpha, \beta)$ to produce d_{ij} , where $j = i$ only on the diagonal.

To produce a single distance value from multiple observations or target pixels, an RMS distance can be used as described in (1.24). A root-sum-squared (RSS) evaluation, used in [9], would imply a target size dependence in pixels on contrast. Thus, larger targets would be given higher distance or contrast values than smaller targets. Though the size dependency rationale may actually be appropriate, an initial size-invariant measure should be established, which can later be scaled in accordance with target size.

RMS Mahalanobis Distance $D(y)$

$$= \left[\frac{1}{n-1} \sum_{i=1}^n \{(y_i - \bar{x})S^{-1}(y_i - \bar{x})^T\} \right]^{1/2} \quad (1.26)$$

Here, once again, y_i is the i^{th} m feature element of n observations, such that $y_i = [y_1, \dots, y_m]$. And \bar{x} represents the mean values of the m features, so that $\bar{x} = [\bar{x}_1, \dots, \bar{x}_m]$. The equation in [9] can be interpreted as calculating the RMS of the Mahalanobis distance. The RMS metric gives greater weight to larger valued elements or distances for consideration. To give an equal or linear weight to each distance measure requires the mean estimation to be moved outside of the square root. The RMS may be appropriate, however, because extremes or larger valued elements can often dominate detection and recognition. This is certainly the case for the task of detection and identification of vehicles in IR imagery.

And thus, a distance measure between a target and background can be determined. The distance or difference between target and background is only a portion of the original notion of contrast expressed in [1]. In [1], small differences in luminous intensity are perceived more when the average luminance of the scene is less. The distance measure in [9] addresses spatial variation and co-variation of pixels from targets and backgrounds, which are not of uniform intensity. To incorporate the distance concepts of [9] and [1], a contrast metric could be defined here as

$$Contrast_7 = \frac{\text{Mahalanobis Distance } D(y)}{\bar{y} + \bar{x}} \quad (1.27)$$

where \bar{y} and \bar{x} are the target and background mean, respectively. The denominator assumes equal presentation of the target and background, but can be weighted to reflect the relative presentation amounts. Here, the average or mean can be defined as

$$\text{Mean}(y) = \bar{y} = \frac{1}{nm} \sum_{i=1}^n \sum_{k=1}^m y_{i,k} \quad (1.28)$$

Multi-Frame Image Analysis

The prior analysis of targets and background discusses techniques which can be applied to target detection within a single image. In many applications, a single image may not be sufficient to describe or detect a target of interest. In some cases, the target of interest is not fully visualized in a single image, and a sequence of images, or video, must be analyzed to generate a suitable signature for both detection and classification. In [10], a linear pyroelectric sensor array tilted at a 45° angle, and later, a sparse pyroelectric array of sensors [11], generates images of target using multiple samples over time. As the target traverses the field of view of the sensor, a portion of the target is captured at each sensor sample and an image is reconstructed of the target using samples from multiple single element pyroelectric detectors. The images were normalized by calculated target velocity estimates, before classification, to allow different target classes to be compared on the same scale for assessing target width.

In another application, presented in [12], image analysis is performed on video imagery affected by turbulence. In this work, spatial degradations in the image vary in time due to turbulence. Occasional glimpses of non-degraded portions of the image are observed over time. These “lucky” glimpses are detected by comparing the edges in the current image to the average of edge measurements observed in previous frames. Multiple lucky image regions can be used to reconstruct a complete image of the scene, where other targets of interest can be detected. Alternately, in [13], regions of strong edges in the current frame are compared to those same regions in the prior frame, to determine image degradation for degraded visual environment scene conditions.

Another more common technique is the use of frame-to-frame differencing. This technique is often used to detect portions of an image that have changed from one image frame to the next, with the assumption that the majority of the scene is stationary. This is often the case for a moving target. The prior frame is treated as a background frame and is subtracted from the current frame to create a difference image. In thermal imagery, targets of interest often have higher intensity than background surroundings. As a result, the current location of the moving target is highlighted in the difference image. The location where the target appeared in the prior image can also create a negative residual in the difference image, known as ghosting. Ghosting can occur both for image subtraction and also image addition, when targets are moving in the scene. And in more general terms, ghosting can occur when the current frame is processed or filtered based on contributions or data from prior frames where motion occurs.

In frame-to-frame differencing, the target is detected in the current image by applying a threshold to the difference image generated from prior image frame data. Another approach, but similar in concept, is to generate statistics from N prior background frames, and use the generated statistics to determine a threshold for determining the target in the current frame. A common approach is to determine a detection if the target intensity is a number of standard deviations from the background mean intensity. This approach can allow adaptable threshold generation for a constantly changing background, such as is observed when the sensor is moving. To generate the desired background statistics, a history of prior frames is often stored and updated in memory via a sliding window, and accessed for each frame update.

In some practical real-time and/or remote systems, storing and processing large amounts of data for background statistics may not be ideal. In this case, it may be beneficial to use a running statistic, such that at each frame, a mean and standard deviation are calculated from the

current frame, and the previously calculated mean and standard deviation from the prior frame [14]. The mean of a pixel of the current frame can be found by

$$Mean_{Running} = \frac{N-1}{N} mean_{prev} + \frac{1}{N} value_{current} \quad (1.29)$$

where $mean_{prev}$ is the mean calculated for the previous frame, $value_{current}$ is the current value of the pixel, and $N-1$ is the number of frames in the window. To determine the standard deviation of a pixel in the current frame, the following equation can be used:

$$\begin{aligned} & standard\ deviation_{Running} \\ & = \left[\frac{N-1}{N} std_{prev}^2 + \frac{1}{N-1} (value_{current} - mean)^2 \right]^{1/2} \end{aligned} \quad (1.30)$$

Here, std_{prev} is the standard deviation calculated at the previous frame. It should be noted, once again, that this method does not use a sliding window. The sliding temporal window allows the algorithm to handle dynamic backgrounds. The running statistics window is ever increasing from the starting location, which causes the statistics to be less susceptible to change as time passes.

To make the running statistics method more accommodating for changing backgrounds, two staggered running statistics can simultaneously be generated for a limited maximum window and repeated. Repeating a single running statistic for a given maximum window can cause noticeable and abrupt changes to the background statistics as the transition occurs, when the window repeats. This is due to the fact that the running statistics needs to evaluate a number of frames before the statistics are representative of the background distribution. By staggering and alternating, one running statistic, $rs1$, can be used for evaluation, while the other running statistic, $rs2$, builds a history. When the maximum window length for $rs1$ is reached, $rs1$ starts a new window. At the same time, $rs2$ is then used for evaluation, while $rs1$ is now building history. This method of staggering and alternating windows improves upon the smoothness of the

transition between one set of background statistics and another. This method can be extended to more than two running statistics. The improved dual staggered and alternating running statistic method would then require 4 frames of memory, the current frame, and the values N_1 and N_2 for the respective windows. In using a running statistic or a sliding window, care must be taken to limit the effects of target signal being added and included in the background statistics, as this can reduce the effectiveness of detecting future targets. In [14], a discussion is provided to address this concern. Once again, the goal of the staggered and alternating running statistics is to continually use relevant statistics, such that resetting the statistics does not introduce significant artifacts within the image or processing. Methods to apply a larger weight to current frame measurements when calculating the mean and standard deviation can be more susceptible to outlier or spike signals.

In many applications, signatures which are detected through the use of multiple frames, are affected by motion. The motion can be due to a moving target, scene motion caused by a moving camera, or both. The presence of scene motion presents an even more complex scenario. One potential technique to reduce observed scene motion is to register image frames. In [15], a survey of image registration methods is provided. For image registration, when scene motion is present, only a portion of the image may contain overlap for all registered images, and may impact detection and processing performance across the full field of view of the sensor.

General Detection Systems

Detection systems, designed for low-cost practical applications, such as remote monitoring have been previously designed. Streetman et al. propose a method and system for detecting events to alert a user [16]. The method is designed to take advantage of input from multiple sensors, which may include more than one domain. The response from multiple sensor

inputs are associated. Data that has been associated from multiple sensors is used to identify events and an event notification is provided to the operator. In theory, this type of sensor data association is used to not only reduce false alarms, but also reduce the cognitive load of the operator. Here, the operator is not required to monitor the outputs of all sensors, but only the determined event notifications. For example, security monitoring of multiple cameras may not be feasible or cost effective, especially in remote sensing applications. A multimodal solution can potentially improve performance of detection and false rejections, while minimizing equipment costs. Streetman, however, provides a general methodology, without identifying a particular system to address a particular problem or application. Although a detailed detection methodology is not provided, a conceptual method of association is outlined, based on identification, location, and temporal pattern. The performance with respect to practical implementation is not assessed in Streetman's work.

Muzzle Flash and Blast Phenomenology

In describing muzzle flash and muzzle blast generation, the following description describes the firing process of small arms, such as the firing of a 7.62 mm NATO rifle. The process, however, can be scaled to describe large arms weapons, such as the firing of 55 caliber artillery. The reader is directed to [17] for a thorough description of muzzle flash and blast phenomenology.

Upon the ignition of the propellant charge in the gun chamber during the gun firing process, pressure and heat build-up, propelling the projectile forward through the gun barrel or tube. The barrel contains air, as well as residual combustion products from previous firings. The air, combustion products, if any, and leaked propellant gases are compressed and forced ahead of the projectile, at high pressure, out of the muzzle of the gun. This expelled air and gas at high

pressure with respect to ambient air pressure, or precursor flow, prior to the departure of the projectile, generates a shock wave, or precursor shock. This muzzle blast or precursor shock radiates, spherically, in all directions, but has the greatest strength along the axial length of the gun barrel.

As the compressed and heated precursor flow is expelled, a bright radiating cone, or Mach cone, is visualized, with the base at the opening of the muzzle in Figure I.13. The expelled gas also interacts with the expanding outer precursor shock wave, such that the gas is contained or restricted by this outer precursor shock wave from expanding. As a result, the gas forms turbulent vortices downrange from the Mach cone, but is constrained by the outer shock precursor wave. As the outer precursor shock wave expands, a cylindrical jet flow region, or shock bottle, is formed, bounded by lateral barrel shock due to the reflecting and coalescing of expansion waves. The jet flow region is terminated by a vertical curved pressure boundary or shock front, called the Mach disk, downstream from the muzzle and Mach cone. The jet flow is comprised of the expelled compressed gases at the Mach cone established at the muzzle, forced into ambient air pressure, creating low pressure, cooled, and supersonic, under-expanded gases. As the gases are propelled downstream, pass the Mach disk and escape the jet flow region, the gas flow is decelerated to subsonic velocity, pressure is increased, and shock heating at the Mach disk causes the temperature to increase by hundreds of Kelvin or more [17]. The increased temperature generates thermal radiation, and in some special cases, a “pre-flash” may occur when turbulent gases containing leaked propellant and combustion products, interact with oxygen in the air, further downstream.

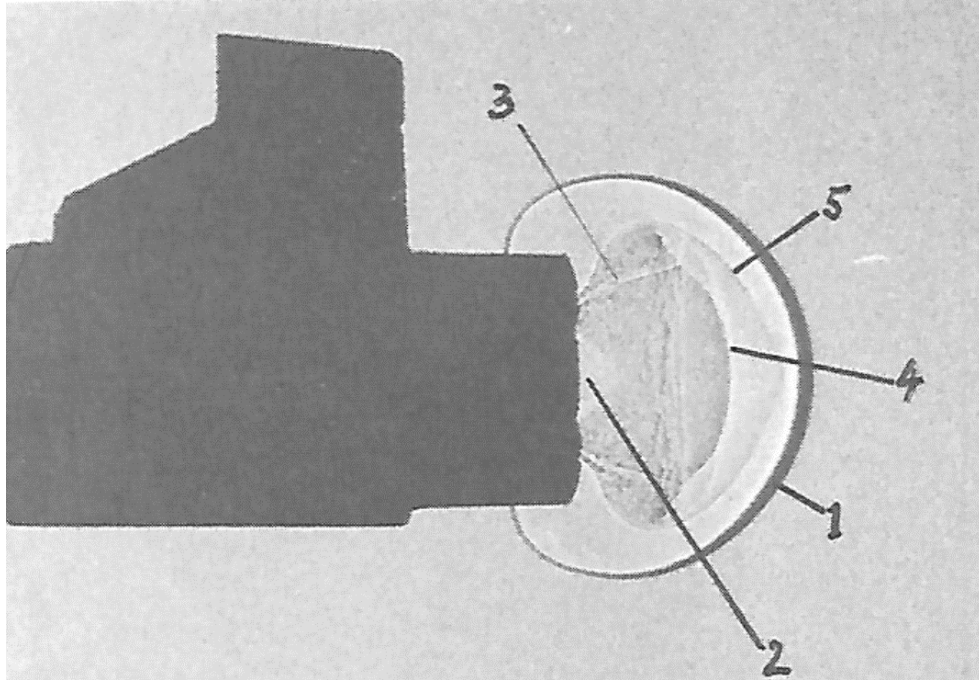


Figure I.13. Shadowgraph of precursor flow. (1: blast wave; 2: Mach cone; 3: barrel shock; 4: spherical flow discontinuity; 5: gas-air interface) [17] G. Klingenberg and J. M. Heimerl, Gun Muzzle Blast and Flash, vol. 139, A. R. Seebass, Ed., Washington DC: American Institute of Aeronautics and Astronautics, 1992, p. 113.

After the precursor flow has been fully formed, the projectile exits the muzzle or tube, followed by the propellant flow gases. A similar process as described for the precursor flow repeats as the propellant flow is ejected. The high heat and even higher pressure of the propellant flow, impeded by the departing projectile, entering into the low-pressure jet flow of the precursor, generates jets or streamers at the muzzle. A blast wave is formed by the propellant outflow, as depicted in Figure I.14. This propellant flow blast wave approaches and overtakes the precursor blast wave as the projectile traverses downrange, exiting the outer blast wave. A similar, yet larger jet flow region and downrange turbulent gas plume are generated. The prior, smaller precursor turbulent gas plume is, in effect, pushed further downrange, such that two gas plumes or gas balls are generated. The precursor jet flow is replaced by the propellant jet flow, as the propellant gases exit the muzzle. For lower velocity projectiles and/or longer barrel weapons,

a second precursor flow and blast wave are allowed to form during the increased time taken for the projectile to exit the muzzle. The effect is an even greater forward the downrange bulge of the gas plume, extending at a greater distance from the muzzle. The gas plume contains the decelerated and heated gases, which are constrained by the blast waves. As the blast wave expands, the heated gas plumes are allowed to mix with ambient air or oxygen.

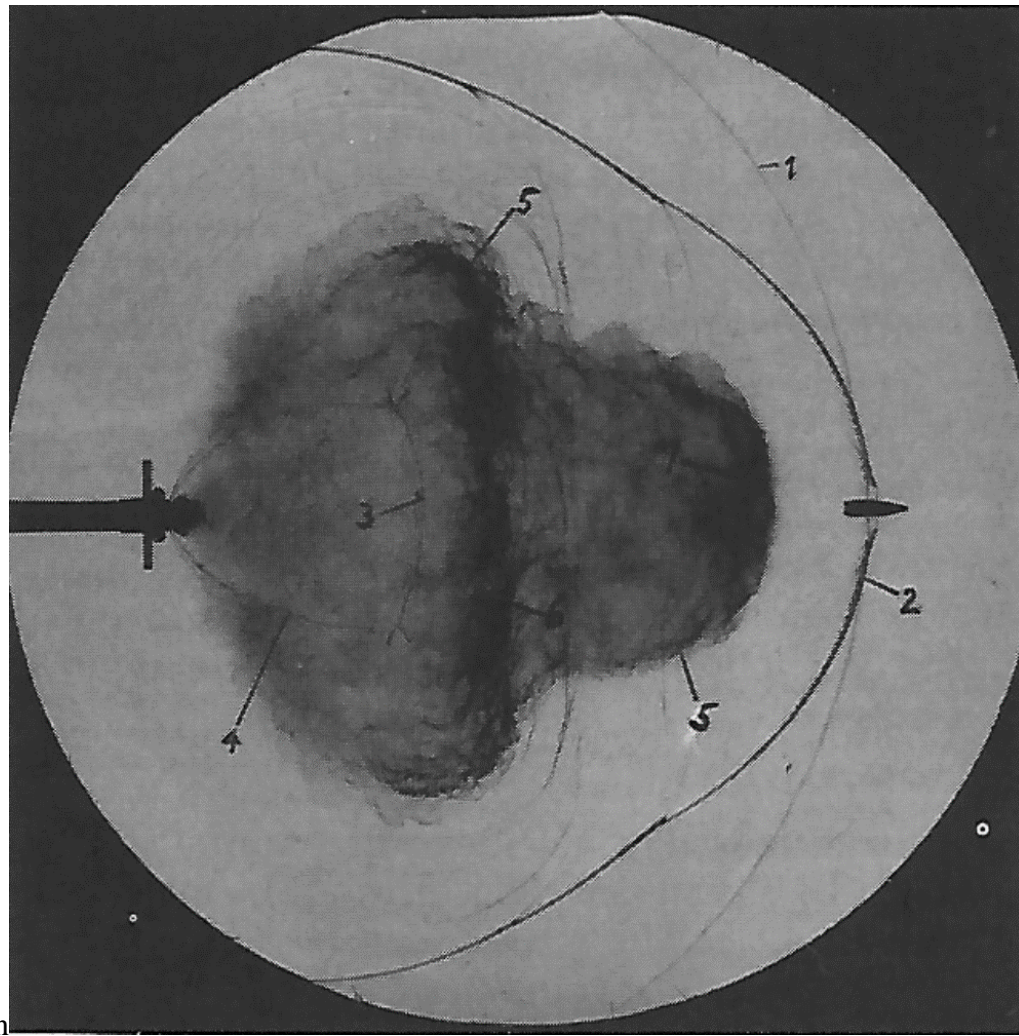


Figure I.14. Shadowgraph of propellant gas plume. (1: blast wave of first precursor flow; 2: blast wave of propellant flow; 3: Mach disk; 4: barrel shock; 5: turbulent gas-air interface; 6: turbulent vortex ring; 7: increased gas density at downstream propellant gas plume) [17] G. Klingenberg and J. M. Heimerl, Gun Muzzle Blast and Flash, vol. 139, A. R. Seebass, Ed., Washington DC: American Institute of Aeronautics and Astronautics, 1992, p. 120.

There are three main flashes observed during a gun firing, known as the primary, intermediate, and secondary or afterburn [17] and are depicted in Figure I.15. The primary flash occurs for a short duration with low intensity and spatial extent, located at the muzzle, once the Mach cone fully develops. The intermediate flash occurs later and further downrange, as propellant gases, rich with fuel, are heated as the gases decelerate beyond the jet flow region and Mach disk. The secondary flash occurs after the intermediate flash and is located both upstream and downstream, having a large spatial extent. The secondary flash is produced as a result of oxygen mixing with the fuel rich propellant gases, and residual combustion products. Ignition can be caused due to shock heating, or even the increased temperatures due to the intermediate flash. The secondary flash has significantly greater intensity and duration than the intermediate flash, and generates a secondary blast wave of greater pressure than the primary blast wave. Similar processes can be found in rocket flames at the exhaust.

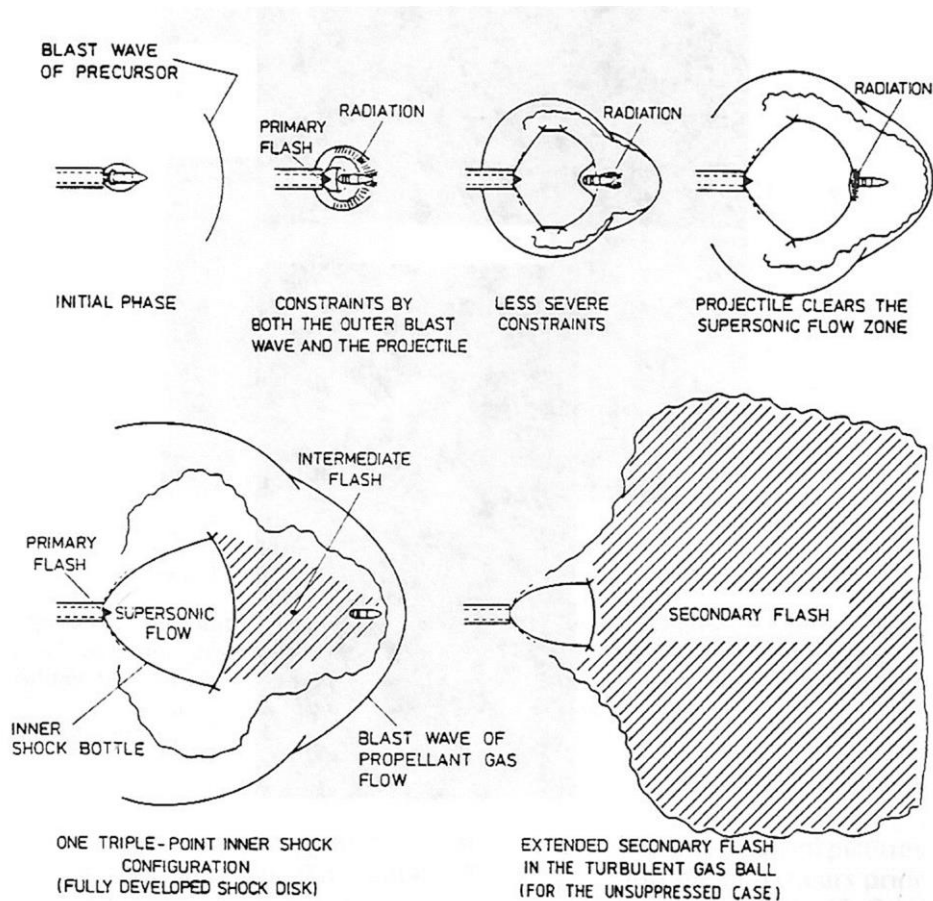


Figure I.15. Development of primary, intermediate, and secondary flashes. [17] G. Klingenberg and J. M. Heimerl, *Gun Muzzle Blast and Flash*, vol. 139, A. R. Seebass, Ed., Washington DC: American Institute of Aeronautics and Astronautics, 1992, p. 214.

The muzzle flash, consisting of primary, intermediate, and secondary flashes, emits radiation across the ultraviolet, visible, and infrared bands. Additionally, the muzzle flow produces a strong continuum of radiation in the visible and near infrared bands due to incandescent flow-borne particles [17]. In general, for blackbody radiation, emitted flux is increased with increasing temperature, and the spectral distribution or peak emittance is shifted towards shorter wavelengths with increasing temperature [18]. Radiation is also emitted along spectral lines due to the chemical composition of the propellant gases and additives. Propellants are composed of carbon (C), hydrogen (H), nitrogen (N), and oxygen (O), such that the ignition

and burning process for both guns and rockets converts the propellants into gases containing oxides of carbon (CO, CO₂) and nitrogen (NO), water (H₂O), and nitrogen (N₂) [17]. Band centers for CO₂ (carbon dioxide), CO (carbon monoxide), and H₂O are 2.349 μm, 2.143 μm, and 3.756 μm, respectively, with CO₂ having the greatest band strength or emitted radiation [19]. These molecules radiate at multiple discrete frequencies. Furthermore, spectral line growth occurs at states of high temperature and high pressure, such that the spectral line increases in emissivity and also in spectral width [19]. The increase in optical path length through the gas, and subsequently the increase in optical depth, can also cause spectral line growth. The emission is attenuated by variations in temperature within the gas, and is also attenuated by lower temperature atmosphere, when observed at a distance away [19]. In [18], graphs of atmospheric transmission for CO₂ and H₂O are provided, showing attenuation notches in the mid-wave infrared and short-wave infrared. The temperature and pressure of the muzzle flash event affects emission and atmospheric transmission of the event signature, and, as a result, the choice in sensing waveband.

The secondary muzzle flash produces a flash of greater extent, duration, and temperature than the intermediate or primary flashes. The chemical nature of propellants is a stoichiometrically imbalance such that there is a lack of oxygen to adequately burn the other propellant chemicals within the chamber. Additionally, a stoichiometrically balanced propellant load may cause damage to the chamber due to pressure. The oxygen deficiency allows optimum energy for propelling the projectile, but as a result, unburned propellant is released in the jet flow. As the propellant jet flow is decelerated, upon exiting the shock bottle at the Mach disk, the propellant gases are increased in temperature. Reduction in pressure boundaries, as the outer shock front expands, allows trapped oxygen in the air to mix with propellant, allowing for the

potential ignition of the secondary flash. Additives are often added to reduce the secondary flash. Common additives shown to be effective, include Potassium (K) and Sodium (Na) [17]. Examples of additive compounds reported to suppress muzzle flash are potassium sulfate (K_2SO_4), potassium nitrate (KNO_3), and sodium bicarbonate ($NaHCO_3$). The addition of flash reducing additives can generate excessive smoke or alter ballistic performance, which is undesirable, and limits the amount of flash reducing additives. Though the additives reduce the secondary flash, radiation is emitted along spectral lines as a result of the added Potassium or Sodium.

Optical Detection Methods and Systems

A weapon fire detection system is proposed by Tidhar, which utilizes uncooled infrared sensors [20] [21]. Observing the increased cost and power consumption of cooled systems, Tidhar presents an optical flash detector system which operates in the near infrared (NIR) and short-wave infrared (SWIR) spectrum. The system is aimed at detecting firearms gunshots, and in particular, sniper and small arms detection. It is argued that the choice of spectral band can be determined to align with atmospheric absorption peaks. Though signal from the muzzle flash will be reduced due to atmospheric absorption, so also will sunlight clutter and shot noise. The purpose is to improve the signal to clutter ratio, assuming the target signal is above the sensitivity threshold. This proposal may, however, limit overall system range performance, which may be less impactful for short range targets. Here, a choice in hardware specifications is used to supplement false alarm rejection processing, by limiting the amount of signal entering the system from false alarm sources.

Tidhar presents options for detection processing in which a layered approach is utilized [20] [21]. In Tidhar's approach, processed pixels found to have signal indicating a possible

detection, can be re-analyzed in a subsequent processing layer, in which the neighborhood of the pixel is also processed. Additional pixel detections found within a neighborhood of a previous detected pixel, can be suppressed, to remove additional pixel processing. This is a form of data management, by limiting the bandwidth at subsequent layers. Tidhar also states that data may be stored in memory using compression or reduced sampling rates. Furthermore, Tidhar suggests a Constant False Alarm Rate (CFAR) module. The CFAR confirms/suppresses a constant predictable number of possible detections at each interval or frame. The CFAR thus bounds the maximum amount of processing, and processing time that will occur at each frame. Thus, each frame is ensured to be processed before the next frame is ready for processing, removing processing delays impacting real time processing. Redundant processing, bandwidth, and real time implementation aspects are addressed by Tidhar.

Multi-Sensor Methodology and Systems

A multi-sensor approach is also discussed by Tidhar. A multi-band approach which employs a sensor operating in the ultra-violet, visible, and NIR spectra is coupled with a sensor operating in the SWIR spectra [22]. It is proposed that a ratio of the shorter and longer wavelength spectrums can be indicative of a muzzle flash. The ratio of direct light from the sun has a significantly different ratio than that of muzzle flashes. Thus, a two-color detection system is employed to reduce false alarms due to solar clutter. Using a simple ratio as a detection method can provide fast and efficient processing for real time applications. For indirect, reflected light, the ratio detection methodology may not maintain effectiveness. For example, when the surface of the reflecting object is rough, with respect to the reflected wavelength, the light is diffused, reducing the amount of light reaching the sensor. When both bands receive light reflected from a surface which is smooth in both bands, or rough in both bands, the relative ratio

between the bands is maintained. In the event that one band experience reflected light from a rough surface, while the other band experience light from a smooth reflected surface, the ratio can be altered enough to generate false alarms. For this reason, the author incorporates a polarizer in the system design to filter out the specular reflection. Here, the author provides a simplified processing methodology, which is found to achieve efficient detection and false alarm rejection under certain conditions. For the conditions in which the simplified ratio method is known to fail, a hardware modification is provided to address and complement the shortcomings of the proposed methodology. Furthermore, Tidhar proposes that an acoustic sensor can be used in conjunction with the proposed optical sensor to provide reduced false alarms, range to flash, weapon identification, and knowledge of environmental conditions [20].

Acoustic Detection Phenomenology

Weapon fire, due to the explosive charge required to propel the projectile from the barrel or tube, exhibit both a visual muzzle flash, but also an acoustic muzzle blast. The formulation of the muzzle blast, propagates a sound energy wave, often referred to as the “bang,” due to the initial explosion of the weapon fire event. Pressure waves, produced by the expelled precursor and later jet flow contribute to form the outer blast wave, depicted in Figure I.16. At longer ranges, the steep, high pressure blast shock wave loses energy and increasingly resembles an acoustic sound wave. The acoustic wave travels in all directions, however, the majority of the energy travels in the pointing direction of the barrel or tube. The wave travels at the speed of sound. The wave reflects against surfaces and interacts with the environment. Examples of environmental factors are temperature and wind gradients, as well as atmospheric absorption [23]. For a microphone located within sensing range, some distance away from the weapon fire event, a response will be generated due to the direct path travelled by the muzzle blast wave.

Later responses may also be generated by reflected waves, due to ground reflection for example, which are delayed due to the longer, indirect travel path of the wave.

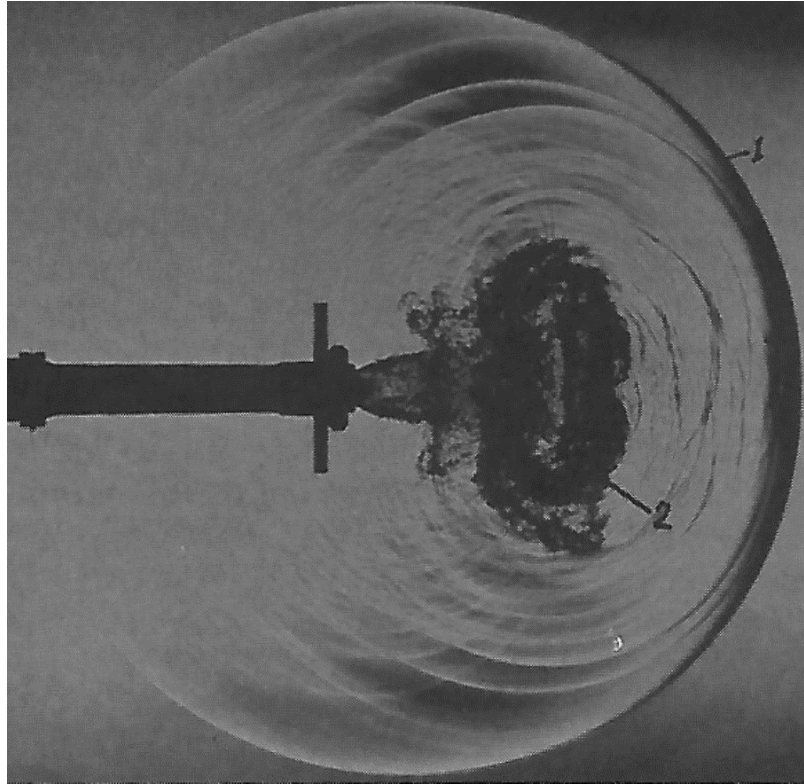


Figure I.16. Dark-schlieren photograph of muzzle flowfield before projectile ejection (1: blast wave of first precursor flow; 2: turbulent vortex ring) [17] G. Klingenberg and J. M. Heimerl, Gun Muzzle Blast and Flash, vol. 139, A. R. Seebass, Ed., Washington DC: American Institute of Aeronautics and Astronautics, 1992, p. 112.

For projectiles travelling at supersonic speed, an acoustic shock wave, often referred to as the “crack,” is generated from the projectile as it travels. The shockwave forms a cone, trailing the projectile. The shockwave propagates outward from the trajectory path of the projectile [24] and travels at the speed of sound. The angle of shockwave propagation is dependent on the speed of the projectile, as depicted in Figure I.17. For a project traveling at supersonic speed, the angle of the formed cone is given by $\theta_M = \arcsin(1/M)$, where M , is the Mach number defined by $M = V/c_s$. Here, V is the velocity of the projectile and c_s is the speed of sound. The speed of sound, in air, increases with temperature, such that

$$c_s = c_0 \sqrt{1 + \frac{T}{273}} \quad (1.31)$$

where T is temperature in degrees Celsius (C) and c_0 is the speed of sound at 0°C (331 m/s) [25].

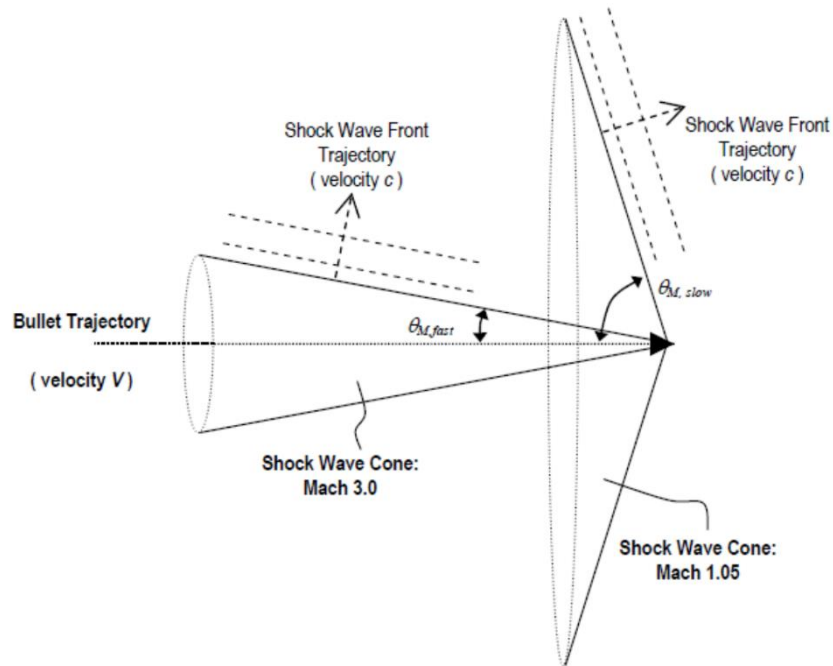


Figure I.17. Shock wave geometry [25] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in IEEE 12th Digital Signal Processing Workshop, Jackson Lake, 2006.

The shockwave propagation is perpendicular to the sides of the formed cone. Thus, for a higher Mach number, a smaller cone angle is formed, and the shockwave's travel path is nearly perpendicular to the trajectory of the projectile. As the projectile slows, the cone angle increases. At the limit, as the bullet becomes subsonic, the cone angle approaches 90 degrees, with a resulting shockwave travelling parallel to the path of the projectile. The shockwave is characterized by an "N" shape due to a fast rise to an overpressure, and subsequent quick decline to an under-pressure. Visually, this shape can be described as a sharp positive impulse followed

by a sharp negative impulse. As the shockwave travels, it loses the sharp “N” shape characteristics and begins to resemble an “S” shape [24] due to spatial spreading. The changing angle of the shockwave, as the speed decreases, causes the energy within the received, “once sharp” impulses, to spread over time. Similar to the acoustic blast, the shockwave will be reflected by surfaces, such as ground reflections (see Figure I.18), and the reflections will be attenuated due to absorption.

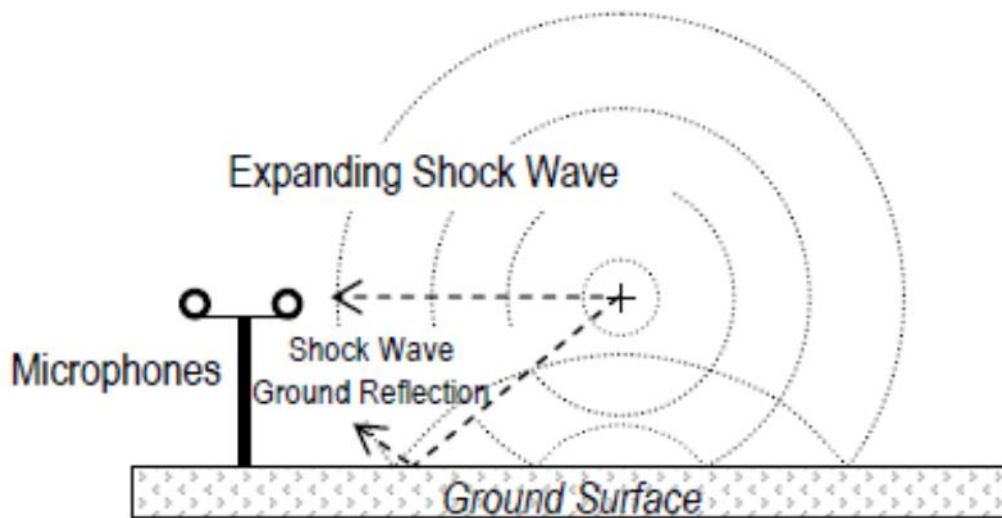


Figure I.18. Shock wave reflections [25] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in IEEE 12th Digital Signal Processing Workshop, Jackson Lake, 2006.

In Figure I.18 two microphones are shown to detect the initial shock wave and the later arriving ground reflections. At a known location, a single microphone can be used to determine the arrival time of the projectile shock wave, the acoustic muzzle blast, and associated reflections. When using two or more microphones at known locations, time difference of arrival techniques can be used to estimate the elevation and the distance of projectile trajectory from the microphones. In Figure I.19, an example acoustic recording, from two microphones, depicts the time difference of arrivals of the shock wave, muzzle blast, and associated reflections. Four or

more microphones can additionally determine the orientation of the trajectory and the projectile speed. The curvature of the shock wave, related to the Mach number, is used to estimate the projectile ballistics at longer ranges, as the projectile velocity decreases due to drag forces. Spatial separation of the detecting microphones is needed to determine the curvature of the shock wave [24]. As an example, a model of a supersonic projectile and associated shock wave and muzzle blast are shown in Figure I.20.

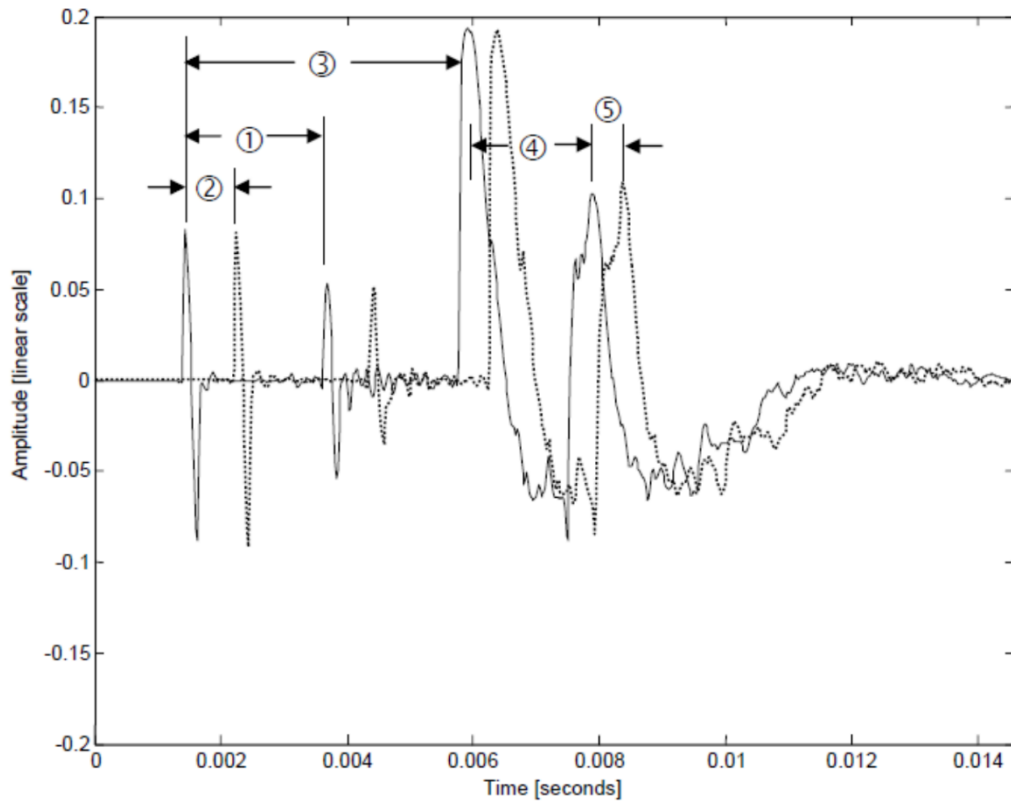


Figure I.19 Gunshot acoustic recording from two microphones emphasizing time difference of arrival of signals (1: shock wave and reflected shockwave; 2: shock wave at near and far microphone; 3: shock wave and muzzle blast; 4: muzzle blast and reflected muzzle blast; 5: reflected muzzle blast at near and far microphones) [24] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in IEEE 12th Digital Signal Processing Workshop, Jackson Lake, 2006.

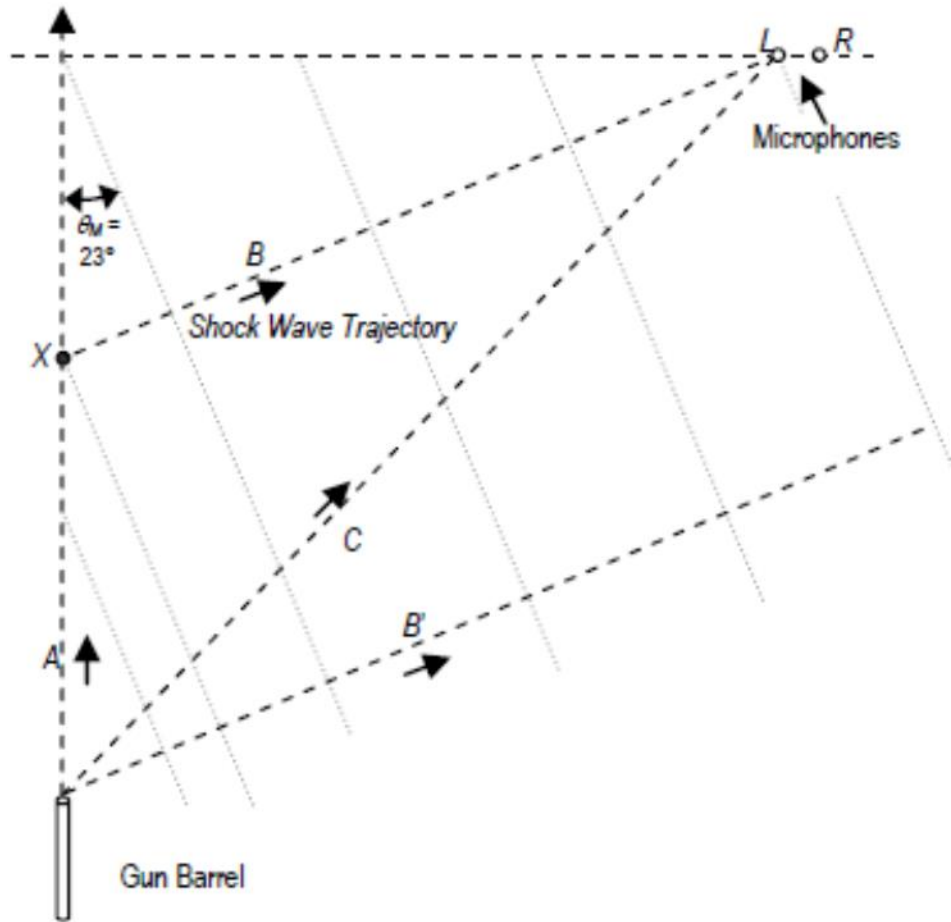


Figure I.20. An example model of supersonic weapon fire geometry [25] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in IEEE 12th Digital Signal Processing Workshop, Jackson Lake, 2006.

Acoustic Detection Methods & Systems

One technique for gunshot detection is to use multilateration (MLAT) [26]. MLAT is a technique used for airport surface surveillance, and has been extended to gunshot detection or surveillance. The technique relies on a distributed sensor network, where the locations of the sensors are known. The emitted signal of the target is received by the sensors. The measured time difference of arrival (TDOA) of these received signals at each sensor location are related to the sensor system location geometry and the unknown target position. The difference in range, from target to sensor locations, is related to time difference of arrival, with the speed of sound,

for acoustic applications. The resulting equation for each pair of sensors defines a hyperboloid representing points of possible solutions to the target location. Theoretically, the target location is determined by the intersection of hyperbolas resulting from detections from sensor pairs [26].

The position accuracy in an MLAT system is affected by three factors: 1) the accuracy of measurements, 2) the spatial locations of the sensors, and 3) the algorithm used to solve for the unknown target location [27]. A non-uniform speed of sound can increase measurement error, along with time measurement sampling. The dispersal of sensor locations can affect the dilution of precision, such that the possible solution space is increased, rather than decreased.

Presumably, the sensors are widely spread, and encompass the target. The number of sensor pairs which are able to detect the target can also affect whether a solution exists [26].

For MLAT systems, the measurement error and spatial distribution of the system, determine the limit of achievable performance. The system is further degraded by the efficiency of the localization algorithm [27]. The localization algorithm can be decomposed into two components: 1) a data model which relates known and unknown parameters via a system of equations and 2) a numerical model which solves the system of equations. Thus, different data models can be applied to different numerical models, and the combination of both should be taken into consideration, when determining algorithm efficiency or performance. In this work [27], the efficiency of the chosen localization algorithm was measured by the numerical efficiency, or the possibility to provide convergence to a solution, and statistical efficiency, the statistical distribution of position error. Here physical system limitations provide boundaries for maximum performance. Additionally, the algorithm allows further degradation to system performance. The algorithm is decomposed into components which can be chosen such that one component complements the other component. The efficiency of the algorithm is determined by

the ability to not “miss” the target, and also the accuracy of the resulting processing of the detected target.

The MLAT methodology relies on known positions of multiple, well distributed sensors and also relies on accurate timing. This can be more easily achieved for stationary systems, such as those embedded into a city network. For mobile, remote, and transient systems, sensor locations may not be well positioned with respect to the target. Thus, it may not be feasible for multiple sensors to encompass the target. Sensors which report the direction of arrival can reduce the multilateration problem to one of a triangulation problem [26]. Furthermore, relative sensor positions may not be known and GPS may not be available in all scenarios and locations.

In [28], acoustical sensors, in the context of intelligence, surveillance, target acquisition, and reconnaissance (ISTAR) are technology instruments capable of performing surveillance and target acquisition, where surveillance is the observation of acoustical components of environmental noise in a systematic and continuous way, while target detection characterizes acoustical signals that contain environmental noise by performing detection, spatial localization, reconnaissance and identification of previously specified targets. Gunshot location systems perform target acquisition through gunshot detection, gunshot origin estimation, and projectile caliber and trajectory estimation. The gunshot detection process is described in [28], as when sensor signals exceed an amplitude threshold, where the threshold can be fixed or variable. Impulse duration, impulse “sharpness” or slope, and directional signal patterns (in the case of multiple sensors), are used to further discriminate a gunshot blast wave from background noise. Frequency content is used to distinguish muzzle blasts against ballistic shockwaves. Gunshot direction of arrival is determined by analyzing the time delay of arrival of the muzzle blast and shockwave, between pairs of sensors to determine direction of arrival. The curvature of the

muzzle blast wave front is analyzed to determine the range to shooter or launch. Projectile caliber can be determined by analyzing the “N” shaped shockwave (N-wave), using empirical relationships between the N-wave peak amplitude, duration, and distance travelled from the projectile. The ballistic trajectory is estimated for supersonic projectiles by analyzing the shockwave arrival time, and the “N” shaped shockwave, across multiple microphones.

In [29], gunshot detection algorithms are evaluated. The evaluation takes into consideration detection performance and realizable hardware solutions, emphasizing power and space. Detection performance is defined by a receiver operating characteristic (ROC) plot which relates true positive rates (TPR) to false positive rates (FPR). The TPR is a ratio of the number of detected true positives to the total number of true positives. The FPR is the number of false positives detected to the total number of false positives. Each point on the ROC plot provides a TPR (y-axis coordinate) and its corresponding FPR (x-axis coordinate), for a given detection threshold value. The Euclidean distance from the measured points on the ROC curve to the perfect detector ($TPR = 1, FPR = 0$) is used to determine detection performance. In [29], short-time Fourier transform, wavelet transforms, hidden Markov models, Gaussian Mixtures, and maximum likelihood models are algorithms noted as having been effective at detecting, classifying, and localizing gunshots, but discusses that these algorithms can require significant power and processing, making them non-viable solutions in remote applications. Instead, algorithms consisting of absolute value, median filter, correlation against template, discrete wavelet transform (DWT), and continuous wavelet transform (CWT) are assessed. Hardware implementations of each algorithm are designed, and the static power requirements are approximated. For comparison, the lower performing absolute value estimator had a Euclidean distance of 0.205, resulting from a TPR of 0.844 and a FPR of 0.133. In contrast, the higher

performing DWT had a Euclidian distance of 0.111 resulting from a TPR of 0.889 and a FPR of 0.0. Though the DWT shows an increased detection performance over the absolute value estimator, the DWT requires 14 single-ended operational transconductance amplifiers (OTA) and 3 rectifiers, resulting in a 3.375 μA bias current. The absolute value estimator, on the other hand, requires a dual OTA, a single-ended OTA, and a comparator, generating a 405 nA bias current. For a given application or set of requirements, such as remote sensing, the choice of algorithm can be a compromise between detection performance, and realizable hardware capable of performing the task under constraints.

Gunshot detection systems have been maturing since the early 1990s after global conflicts led to the development of the Joint Counter Sniper Program by the U.S. Army in 1993 [28]. A Defense Advanced Research Projects Agency (DARPA) initiative, starting in 1995, advanced sniper technologies and led to the development of a number of sniper detection systems. These systems include the Fast InfraRed Sniper Tracker, developed by Thermo Trex Corporation, which detects the muzzle flash in the infrared. Uncooled infrared and acoustic sensors were combined in the Integrated Sniper Location System developed by SenTech, Inc. and Lockheed Martin. Another system, the Sentinel, developed by Science Applications International Corporation, incorporates 16 acoustic sensors. Bullet Ears counter sniper system was developed by BBN Technologies of Massachusetts which incorporated advanced algorithms. The System for the Effective Control of Urban Environment Security (SECURES), developed by Alliant Techsystems Inc. was deployed at the Atlanta Olympics in 1996 and was used in urban cities for gunshot detection. During the same period, another gunshot detection system, ShotSpotter, developed by Trilon Technologies, was also deployed in urban environments.

Position, Navigation, and Timing

The most widely used and the most accurate source of position, navigation, and timing (PNT) data is provided through the Global Positioning Systems (GPS). GPS is provided by the United States, maintained by the U.S. Airforce, and is one of many Global Navigation Satellite System (GNSS). GPS Standard Positioning Service (SPS) delivers continuous PNT information to the worldwide community, free of direct user fees, for peaceful civil, commercial, and scientific uses [30]. GPS, as other GNSS, consists of a ground control station system, a satellite system, and user GPS receiver equipment. GPS receivers accept radio signals transmitted from the GPS satellites and triangulate the receiver's position to within several meters or better, depending on the receiver and environmental factors. As an example, smartphones with GPS capability are typically accurate to within 4.9m under open sky [31]. Additionally, a Master Clock system, through the use of an ensemble of more than 100 atomic clocks, delivers Coordinated Universal Time (UTC) as an underlying timing reference to GPS. The provided timing reference enables users to synchronize to UTC time to less than 40 nanoseconds [31] or within 100 billionths of a second with an accurate frequency reference, without the need for direct access to an atomic clock [30][32].

In addition to GPS, inertial navigation systems (INS), also referred to as inertial measurement units (IMU), are used to compute position and navigation data. The INS is able to provide 15 degrees-of-freedom measurements, which include three each of position, velocity, acceleration, attitude and rotation rate measurements. Performance characteristics of INS are well characterized and consistent, making them reliable. Additionally, the INS is a self-contained system, given initial conditions, and is not subject to external influences. Typical measurement output rates are greater than 100 Hz, providing high frequency navigation updates. Without

periodic updates, however, the INS will drift over time as increased error or uncertainty is included in the measurement process.

The INS contains a computer processor and sensors, to include accelerometers and gyroscopes, to determine position, orientation, and velocity through dead reckoning. Accelerometers are used to measure linear acceleration, while gyroscopes are used to measure angular velocity with respect to an inertial reference frame. Dead reckoning refers to estimating the current position using knowledge of a previous position, and estimates in changes in orientation and changes in velocity, over time. Thus, dead reckoning does not measure an absolute position or velocity, but estimates the position and velocity, based on current sensor data inputs and initial conditions. The accelerometer measures acceleration due to both non-gravitational forces and also gravitational force. The acceleration due to gravitational force must therefore be subtracted from the accelerometer measurement, requiring a gravitational model and knowledge of the reference frame. The gyroscope, measuring rotation about an inertial reference frame, is used to map the measured acceleration to a reference axes. Errors in the gravity model, initial conditions, and numerical computations, as well as sensor errors, are perpetuated to the resulting solution. In this method of dead reckoning, errors such as gyro drift rate and initial heading error cause an increasing INS error over time and greater uncertainty of future position estimates.

When combined with GPS, the INS error can be periodically corrected by providing initial position and periodic position corrections. GPS navigation measurements are combined with INS navigation measurements through a Kalman filter. The Kalman filter estimates the current navigation state measurements based on a weighted contribution of sensor measurements and the prior navigation state measurements. The weights are determined based on estimated

error, including measurement sensor error, such that a measurement with less error is given greater weight. The integrated navigation solution is intended to be more accurate than a solution derived from any individual input sensor. GPS alone does not provide roll, pitch, or yaw, nor does an INS provide an absolute attitude. These navigation updates can be provided through an integrated INS/GPS system. Commercial automotive navigation systems use a combination of GPS and dead reckoning through multiple sensors for increased reliability when GPS signal is not available. The INS is able to generate position and navigation data, in between GPS updates, when GPS updates are not available, and can be used without external reference or sensors, once initialized.

Alternate Applications

Impulse detection can be used for a variety of applications. Though weapon fire detection is a very relevant application in today's society, other applications, such as brain spike detection exist and follow similar methodologies. In [33], spike detection algorithms are evaluated for use in a brain-machine interface (BMI). The brain-machine interface consists of an application-specific integrated circuit (ASIC), which is implanted into the brain. The ASIC detects action potential (AP) waveforms, or spikes. The detected signals and corresponding detection times are transmitted wirelessly to a receiver, which processes the signals to control a prosthesis. In this application, real-time data processing and transmitted data bandwidth are of concern, in addition to detection rates and false alarm rates. In this application, missed detections are more costly than false alarms. Though processing at the ASIC allows for only needing to transmit the detected waveform and detection time, increased false alarms can consume bandwidth as well as processing time. Additionally, the BMI has limited processing capability. In this study, it was found that although matched filters using convolutions performed well, they also required a high

computation time. A combination of simpler techniques, which thresholded both the rising and falling peaks of the signals of interest, proved most effective when evaluating detection rates in conjunction with bandwidth and processing time. Here, a filter was applied such that a refractory period occurred, so that signals which occurred too close in time were rejected. The combination of these simple techniques allows for a remote detection system with limited processing and bandwidth, to relay detected events to a supplemental system for further processing and automation.

In other literature, algorithms and architectures for automatic brain spike detection and alignment, which have been designed for low power, are evaluated [34]. Similarly, to reduce the need for a large number of wires, which can constrain the subject, it is preferable to process the data and transmit the results wirelessly. In this paper by Zviagintsev [34], multiple detection and alignment methodologies are evaluated with respect to their impact on the subsequent sorting operation. Thus, the detection and alignment methodologies are not evaluated in isolation but are evaluated based on their impact to the system. In general, due to unstable recording conditions, spike processing must be adaptable. Raw data is therefore periodically transmitted to the processing computer for re-training and re-calculation of parameters. Thus, the system is able to adapt to a changing processing environment. The spike detection and alignment process, prior to sorting, consists of detecting the spikes in the input and determining their starting point. In one evaluated algorithm, Maximum Projection Alignment (MPA) determines the spike's temporal location by finding the highest correlation between the sample segment and the first principal component, determined from prior data. Upon estimating the spike location, correlation with the second principal component is only then calculated and both results are sent to later processing stages. In another evaluated algorithm, Maximum Integral Transform Alignment (MITA), the

energy under the curve is measured across different segments of the curve and compared to prior data. In this case, two segments are used which distinctly define the characteristics of the signature in question. In a third evaluated algorithm, Segmented PC, the principal components are downsampled, and the sample is integrated into k segments to reduce the amount of total multiplications needed for processing. The algorithms were evaluated by both system classification error and computational complexity.

Organization of this Dissertation

The dissertation is organized in the following way:

Chapter 2 discusses preliminary investigations, motivation, and results in the algorithm development process, focused on weapon fire detection. Detection is demonstrated at various wavebands and video frame rates, and sources of clutter are analyzed. Chapter 3 continues the algorithm development process with increased focus on false alarm mitigation strategies. Chapter 4 consolidates the algorithm, described in Chapter 1 and chapter 2, and presents a system in which the algorithm operates. Chapter 4 also presents alternative systems and methods for mitigating the effects of saturations. The content in Chapters 4 has been submitted and accepted as patents to the United States Patent and Trademark Office (USPTO) [36][37] and is written in the style of a patent submission. Additionally, Chapter 4 provides a multi-modal system design and methodology, incorporating an acoustic sensor to reduce false alarms. The multi-modal design, in Chapter 4, has been submitted as a patent to the USPTO and is also written in the style of a patent. The work defined in Chapters 2, 3, and 4 contributed to the creation of two journal papers submitted to JDRE [38][39]. Furthermore, Chapter 4 provides a methodology, incorporating PNT data to mitigate saturation clutter through sun detection and tracking, and additionally, is written in the style of a patent.

II. DETECTION AND CLASSIFICATION METHODOLOGIES

Introduction

This chapter follows the initial investigation into the feasibility of using temporal profiles for weapon fire detection by using an EO/IR sensor. The work discusses the metrics identified for the detection process and acknowledges the significant issue of clutter for an imaging sensor, for on-the-move application. In addition, results of weapon fire classification are shown, providing the potential for additional capability.

Problem Description

An investigation into feasibility of an EO/IR impulse detection algorithm for weapon fire detection. Provides a baseline methodology for detection, and demonstrates the possibility of weapon classification.

Methods

The Hostile Fire Detection (HFD) algorithm presented here is decomposed into three stages: pre-detection (see Figure II.1), detection (see Figure II.2), and classification. The pre-detection stage distinguishes pixels with signals, which vary from the background. These pixels are then tracked over multiple frames for consistent signal. Pixels which persist for a given number of frames are noted for pre-detection. Given knowledge of actual shot duration times, pre-detects are tracked for a period of time in which an actual shot would have occurred. After this elapsed time, the pre-detections are evaluated to find the pre-detections with a local maximum value in space and time. The local maximum pre-detection is then passed on to the detection stage for further analysis.

The isolated pixel is then assessed over a history of frames to ascertain when the shot occurred. The shot signature is thus extracted for temporal shape profile measurements. The

duration of the shot signature is measured. Additionally, the shot signature is normalized in both intensity and time such that the shot is integrated into 10 discrete samples in time, which add up to unity in intensity. The relative location of the peak intensity, energy distribution under the curve, rise slope, and fall slope are measured from the normalized shot signature. Shot signatures which fall within given thresholds for these measured parameters are considered as valid detections of hostile fire signatures and passed on for classification.

Measured parameters previously calculated in the detection stage are reused as features for classification. In particular, signal duration and shape, provided by the normalized shot signature, are used for classification. Signatures are classified into three broad categories of threat type: CLASS 1, CLASS 2, and CLASS 3. Though not named in this work, CLASS 1, CLASS 2, and CLASS 3 represent three distinct categories or types of threat weapons.

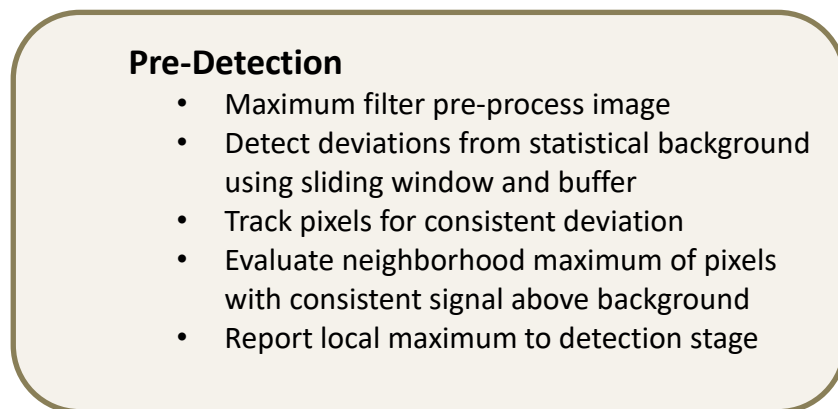


Figure II.1. Quick reference guide for the pre-detection scheme used.

Detection

- Extract shot signature (start / stop time)
 - Deviation from background, change in slope at rise, signal return to percentage of peak
- Measure signature duration
 - Width at 10% of peak
- Measure shape
 - Temporally / intensity normalize signature into 10 samples
 - Peak location, decreasing tail, energy distribution, rise slope, decreasing Δ slope
- Measure amplitude
 - Change in digital numbers (DN) from background
- Measure neighboring pixels
 - Neighborhood summation

Figure II.2. Quick reference guide for the detection scheme used.

Pre-Detections

As stated above, the pre-detection stage is the first of three stages in which a pixel is identified as having consistent signal. In determining signal, a background model is developed, and a signal-to-background ratio (SNR_B) is computed. For each pixel, a temporally sliding window is used to establish mean (μ_{bkg}) and standard deviation (σ_{bkg}) statistics of the background. The sliding statistical background window (in history) is separated from the present, current frame's pixel, by a temporal buffer. The separation provided by the buffer allows for both fast and slow changing signals to be sensed for elevated SNR_B , where SNR_B is defined as

$$SNR_B = (pixel_{current} - \mu_{bkg}) / \sigma_{bkg} \quad (2.1)$$

Clearly, pixels can exhibit high SNR_B when the value of the current pixel is drastically dissimilar from the history of background. This, however, can occur quite often when there is movement in the scene or apparent motion due to the movement of a camera system which is sensitive enough to detect changes in the background.

Scene motion can be addressed through multiple approaches. One solution to minimize scene motion is to use image registration techniques, but this method can potentially be computationally expensive. Alternatively, a smoothing filter can be used to diminish the effects of sharp edge transitions from moving objects due to real motion or parallax. A smoothing filter could potentially reduce actual hostile fire signals. Instead, to reduce the number of false pre-detections evaluated, the frames used to produce background statistics are first spatially filtered for local maximums. The maximum filtering can be accomplished by a sliding 2-D window, which returns a local neighborhood maximum at each pixel. The result of the filtering process is a reduction in scene variability, even during motion. The filter can potentially be applied to the current frame without loss of peak signal from hostile fire events. By applying the filter on the current frame, hostile fire events can more easily be tracked for consistency across multiple pixels. In addition, more hostile fire target pixels will be sensed, but clusters of false alarms may be generated as well. Clusters of pre-tracks, whether signals from hostile fire or false alarms, are addressed later in the pre-detection stage. A moving sensor platform, range of targets, instantaneous field of view (IFOV), frame rate, and duration of hostile fire event are factors to consider for applying the maximum filter on the current frame. When hostile fire motion is not significant within the period of regard, simply sensing the peak pixel is adequate for detection and improves efficiency.

At each frame, each pixel's SNR_B is assessed as to whether it exceeds a given threshold. Pixels which exceed the given SNR_B threshold for a consecutive number of frames are noted as pre-tracks. If indeed, the pixel corresponds to a hostile fire event, then a finite amount of time is needed for the hostile fire signature to develop and decay. Given prior knowledge of hostile fire duration times from field collected data, pre-track pixels are investigated at a set m-number of

frames in the future, as shown in Figure II.3. This of course introduces a set amount of latency into the process. An alternate method would be to dynamically evaluate the hostile fire signature as it develops. The dynamic evaluation was not pursued in this investigatory study.

After the hostile fire signature has been allowed to mature, only pre-track pixels are reviewed. All pre-track pixels within a spatial neighborhood of an arbitrary pre-track pixel are evaluated in the m-number of frames in history which were required for the hostile fire signature to develop. The pixel with maximum value is returned and the rest are discarded. The intent here is to extract the pixel with the peak signal value for further processing to determine shot detection, and limit processing on redundant information. Once a detection is made, neighboring pixels can be considered for further discrimination or performing classification. The pixel found to contain the peak signal, deemed a pre-detection, is then supplied to the detection stage.

Depending on system parameters, such as spectral bandwidth and sensitivity, and also scene content and motion, a large number of pre-detections can possibly occur. These pre-detections are then evaluated in the detection phase. To minimize the number of required evaluations, one observation can be made concerning the frequency in which pre-detections occur for a single pixel. For a real hostile fire event, as the pixel is tracked, the sliding statistical background window shifts across the shot signature. During this time, it is unlikely a new event will be detected due to the significant background statistics generated by the hostile fire event. Given expectations of minimal durations of hostile fire events, a minimal duration can be set between pre-tracks on the same pixel. Therefore, if a second pre-track occurs under the minimal duration between pre-tracks, the first pre-track can be immediately discarded without further analysis. This will both remove unnecessary evaluations and also minimize the number of pre-tracks maintained for each pixel.

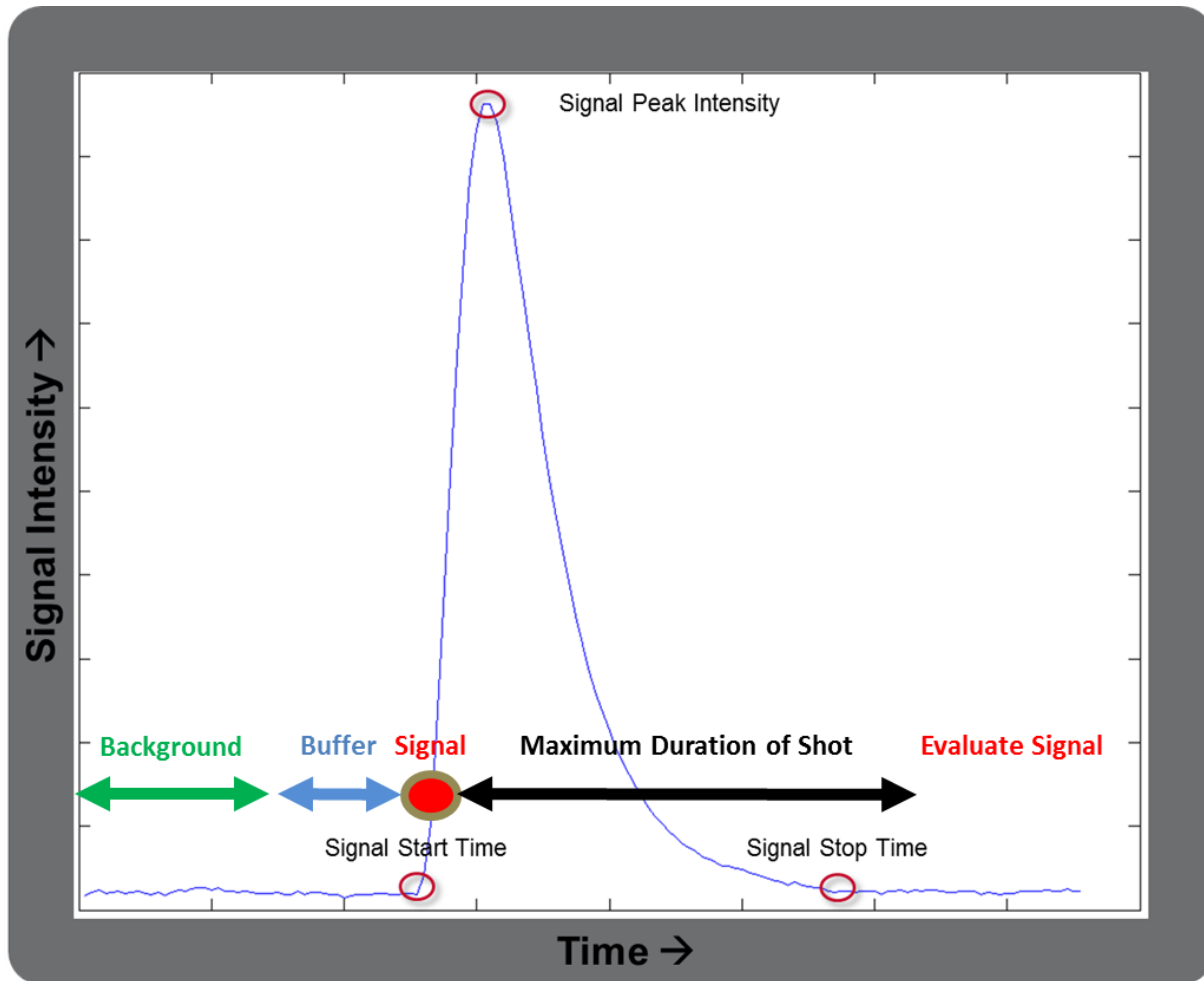


Figure II.3. An illustrative example depiction of the pre-detection evaluation process, which assesses pixels that deviate from the statistical background after a period of time.

Detection

At the detection stage, the potential hostile fire signature is extracted from the pre-detection. Recall that a pre-detection is a pixel which varies from the background and has the peak intensity within a local (spatial/temporal) region. For a real, physical, hostile fire event with known peak, there exists a limited time in which energy is injected into the system, forming the rise time, and a duration in which a significant amount of the energy escapes the system, or fall time. Based on measurements from real data, a reasonable temporal window can be defined, surrounding the pre-detection at the location of peak intensity, in which there can be expectations

for the full hostile fire signature as having occurred. Thus, the hostile fire signature is extracted from the single pre-detection pixel within these temporal bounds.

The extracted hostile fire signature is the basis for performing additional measurements used in discrimination and classification. For a single pixel, a start and stop time establish the hostile fire signature. Variations in the start and stop times, however, can potentially yield variations in additional measurements. It was found that using the statistical background window and buffer system, as described previously, to determine variation from the background, was adequate in most cases for determining a start time. The fast rise in signature, when a true hostile fire event occurs, will be detected immediately as varying from the background. In some cases, though, noise will produce a premature start time measurement. Observing that the signal intensity of hostile fire events suppress signal noise, hostile fire signatures typically exhibit a positive slope until the peak is reached. Investigating the slope of the signal at the start time for negative values can potentially delineate between background noise and hostile fire signature, and give greater confidence in start time measurements. The stop time can also be evaluated with a statistical background window and buffer system. In this case, the background is established post-shot, and the signal is retraced to find variations, which returns the stop time. Once again, greater confidence can be placed in this measurement by analyzing the tail of the signature. The exponential decay of the signature can allow for noise to be captured in the tail of the signature as it approaches the background. A simple slope check may not be suitable. Another method is to determine when the signal intensity returns to a percentage of the peak intensity. Additionally, the standard deviation of the signature can be investigated to disclose when the signature standard deviation returns to a value similar to the pre-shot background. These additional

measures can give greater confidence that only the hostile fire signature is extracted completely, and consistently, for every shot.

For an extracted signature, temporal duration and shape features are measured to verify the detection of an actual hostile fire event. Initial analysis showed temporal duration to be drastically different between noise spikes, small arms hostile fire, and large arms hostile fire, and so, became subject to inquiry. Multiple techniques for measuring signature duration were explored. Clearly, the start and stop times provide a measure of signature duration. An alternate method is to measure the width of the signature at a percentage of the peak. Analysis was performed measuring widths at 10% and 50% of the signature peak (see Figure II.4). Another method is to measure the duration needed to reach 95.45% of the area under the curve. Additionally, treating the signature as a Gaussian distribution, and defining four standard deviations as the signal duration was considered. For a Gaussian distribution, 95.45% of the energy is contained under the curve within \pm two standard deviations from the mean. Similarly, treating the signature as an uneven distribution also led to a measure of signature duration. Moreover, the related measures of rise time and fall time were analyzed.

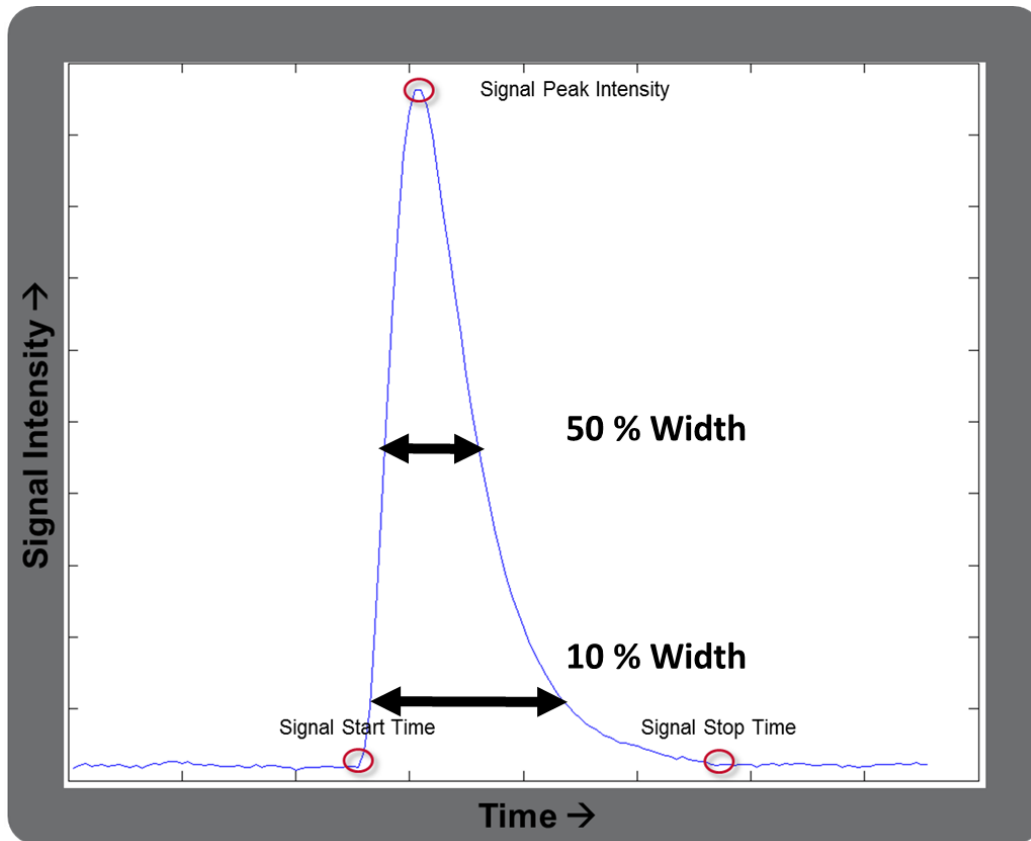


Figure II.4. An illustrative example of a hostile fire signature depicting the width measured at 10% and 50%.

The signature duration can be used as a primary discriminator for detection, but to minimize false detections, further signature analysis is required. Signature features which characterized shape qualities were assessed. Signature intensity and duration can vary dramatically across and within threat classes, yet still maintain a distinct shape or distribution typical of hostile fire signatures. In order to expand false alarm rejections, as well as compare shape features from different classes on the same scale, the signature was integrated temporally into a standard number of samples and then normalized in intensity (see Figure II.5). The number of samples was arbitrarily chosen at 10. It is recommended that the chosen number of samples be small enough to eliminate the effects of noise and irregularities in the signature, while still maintaining observed trends in shape features. The data reduction can also allow for the samples

themselves to be used as features for discrimination and classification. In analyzing the normalized signatures, it was found that actual hostile fire event signatures exhibit a peak at consistent locations, contain more than 70% of their energy within the first half of the signature, experience the greatest change in energy at reliable intervals, and decrease monotonically after the peak has occurred. In addition, signature tails have an exponential decay, such that the change in signal monotonically decreases. It should be noted that these measures demonstrate consistency when initial signature start and stop times are consistent. Thus, recorded signatures which extend beyond the actual hostile fire event, or perhaps, exclude a portion of the hostile fire event, may not display the same regularities.

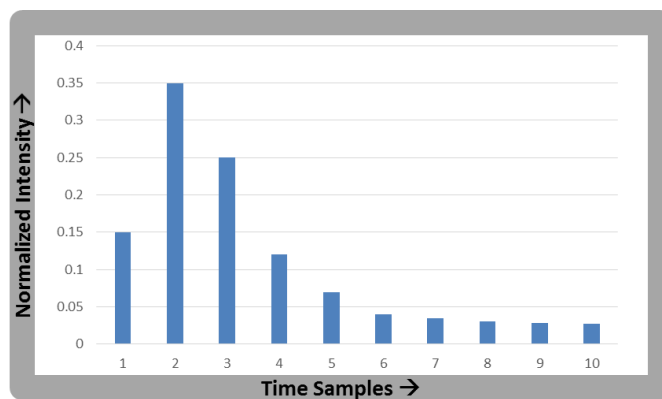
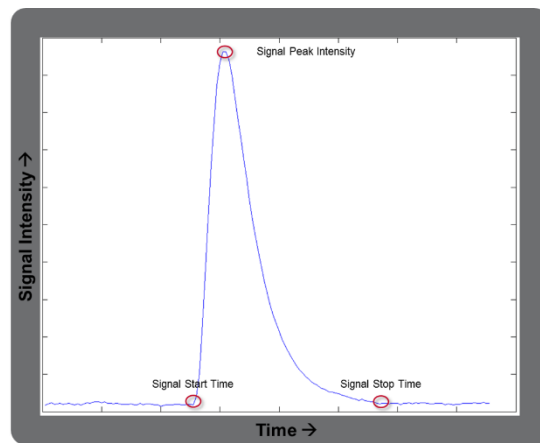


Figure II.5. (Above) A hostile fire signature. (Below) An illustrative example depicting the concept of temporally and intensity normalized signatures (above).

The methodology for detection presented thus far requires only a single pixel on target and is system independent. The hostile fire event need only be observed in a single pixel for a detection to occur. In real data, however, strong intensity of hostile fire events and optical blur can lead to a hostile fire signal being observed in multiple pixels, even when the event is sub-pixel. This phenomenon can be capitalized upon by using multiple pixels in the detection process to give greater confidence that a true hostile fire event has occurred. One approach to reject noise is to investigate the covariance and correlation of neighboring pixels using a sliding temporal window. This method can be computationally expensive, however. Another approach is to median filter pixels within a local spatial neighborhood. Neighboring pixels which detect the hostile fire event will exhibit temporal signatures similar in shape and duration. Median filtering will extract a local signature if more than 50% of the pixels within the neighborhood contain signatures. A summation can also be used, but in this case, a single extreme pixel can dominate the results. For a real hostile fire event, the result of the median filter and the result of the summing filter will also generate a comparable signature. In the former, because each frame is median filtered spatially, noise can be introduced in the resultant signal. A more robust approach is to use a combination of the two methods and compute an average of values near the median value. This has the effect of eliminating outliers and noise, while retaining the general shape of the hostile fire signature. Shape features can then be extracted, however, these features may not coincide exactly with values measured from the maximum pixel and an alternate set of constraints may be more appropriate. The strong intensity of hostile fire events can also be potentially exploited by modeling the sensor system and predicting the expected digital number counts produced. These predictions can be used to exclude noise and clutter which exhibit similar shape as hostile fire, but not the anticipated signal amplitude. It should be remarked that

the intensity rejection approach can possibly limit the system's ability to detect non-ideal threats. Threats which are out-of-focus, attenuated due to atmosphere and range, or are possibly obscured, may exhibit significantly reduced signal. These are possibilities which should also be accounted for.

Classification

In an early analysis of the data, certain features were found to have greater reliability and less variation. Initially, multiple methods of measuring signal duration were investigated. The width at 10% peak height, the width at 50% peak height, the width at 95.45% of the area under the curve, the width as measured by four standard deviations, and also the width measured as four standard deviations treating the signature as an uneven distribution. Though all measured a real phenomenon and yielded similar values, the 10% width and 50% width provided the least variation and greatest separation between threat classes. The separation between any two threat classes was measured in units of RMS standard deviations defined as

$$\text{RMS standard deviation} = \sqrt{0.5 * (std_1^2 + std_2^2)} \quad (2.2)$$

where std_i is the standard deviation of class i . The classes also showed the greatest degree of separation in the LWIR data. Additionally, temporally/intensity normalized signatures and signature fall time duration also demonstrated less variability in the LWIR. With the exception of the 50% width, these features are calculated in the detection phase. In a preliminary analysis, the 10% width, the 50% width, fall time, and the 10 temporally/intensity normalized samples were used as features for classification.

Results

Detection

The hostile fire detection algorithm was tested against real data of threats consisting of CLASS 1, CLASS 2, and CLASS 3 hostile fire. Data was collected for these threats over the course of four data collections. The environments consisted of freezing temperature and snow, warm desert, and moderate green fields with tree lines. Each threat class appeared across at least two of the three environments. Within each threat class, there existed at least two variations of each threat type. In total, approximately 116 hostile fire events were used in this analysis. The probability of detecting hostile fire events using data collected from MWIR-1, MWIR-2, and LWIR 600 Hz imagers is depicted in Table II.1.

Table II.1. Probability of detection for hostile fire events using data collected from MWIR-1, MWIR-2, and LWIR imagers.

Original Data 600 Hz			
Node	Total	Detects	Prob Det
<i>MWIR-1</i>	115	115	100.00
<i>MWIR-2</i>	116	116	100.00
<i>LWIR</i>	116	116	100.00

The performance of the hostile fire detection algorithm is not only measured in its ability to detect hostile fire events, but also in its ability to reject false alarms due to noise and clutter. For this reason, data was collected of potential false alarm clutter sources. The data was collected on-the-move from scenes containing sun glints on cars in parking lots, sun glints on the Potomac River, direct sunlight through trees, sodium lights, moving car headlights, 60 Hz welding, and a burning barrel and flares. The data was collected at speeds ranging from 5-15 mph. In total, 107

videos were collected, each of which was 101 seconds in duration. The false alarm data was collected over three days, one of which was cloudy, while two were sunny.

Table II.2 shows a breakdown of the number of false alarms generated on the cloudy day as a result of each layer of discrimination in the detection scheme, and also as a result of clutter environment type. For the detection scheme, the number of pre-detections is shown, as well as the number of pre-detections after the max window is applied. Each column thereafter refers to a successive layer of discrimination based on 10% peak-width, the peak location, monotonically decreasing tail, energy distribution of signature, and slope of signature rise, respectively. It is clear that application of the maximum filter window significantly reduces the number of pre-detections evaluated. Also, LWIR imagery generates significantly more false alarms. This is due to terrestrial objects at approximately 300 K having a peak emittance in the long wave spectrum, but having significantly lower emittance in the mid-wave spectrum. For a sub-band filtered MWIR imager, with a high noise equivalent temperature difference (NETD), these objects are not observed, but in the LWIR, moving terrestrial objects can be a source of false alarms. Recall that pre-detections are essentially locations which deviate from the statistical background. In a moving scene, with moving objects and parallax causing occlusion, there can be considerable deviation from the current scene and established statistical backgrounds. In Table II.3, false alarms are generated as a result of processing all of the cloudy and sunny days. It can be observed that MWIR images have a significant increase in false alarms due to solar reflection.

Table II.2. On-the-move false alarm results for a cloudy day. (Top) False alarms after each layer of filtering. (Bottom) False alarms per clutter environment.

Node	Pre-Detections	Max Window	10% Peak Width	Peak Location	Decreasing Tail	Energy Distribution	Rise Slope
<i>MWIR-1</i>	7,145,484	384,452	18,869	6,593	299	150	97
<i>MWIR-2</i>	7,600,653	522,733	75,898	25,273	445	132	19
<i>LWIR</i>	76,295,818	29,953,062	3,034,770	932,226	242,949	122,552	50,831

Node	Water	Car Glint	Sodium Lights	Sun Through Trees
<i>MWIR-1</i>	0	0	0	97
<i>MWIR-2</i>	2	1	0	16
<i>LWIR</i>	14,149	29,297	6,405	980

Table II.3. On-the-move false alarms across cloudy and sunny days per clutter environment. General trends in false alarm rates per spectral band and across environments can be observed.

Node	Water	Car Glint	Sodium Lights	Sun Through Trees	Head Lights	Burning Barrel	Welding	Total
<i>MWIR-1</i>	464	3403	0	225	50	105	0	4247
<i>MWIR-2</i>	29	506	0	535	1	14	1	1086
<i>LWIR</i>	35996	67050	6405	2452	8980	746	3	121632

Initial results of probability of detection and associated false alarm rates provide a baseline measure of algorithm performance. With 100% probability of detection, a favored approach is to restrict the detection process to reduce false alarms, while maintaining acceptable probability of detection levels. There is a tradeoff, however, with restricting the detection process and testing on an incomplete data set. The detection scheme should be able to accept new data, as additional data is collected for testing. Beyond algorithmic tradeoffs, there are system level

tradeoffs that can affect the performance of the algorithm, such as frame rate, IFOV, and noise equivalent irradiance.

The effects of frame rate reduction on algorithm performance were analyzed. The frame rate was reduced by temporally downsampling the original 600 Hz data. The data collection imager samples pixels at a significantly higher rate than the readout rate of 600 Hz. For example, removing every other frame, yielding 300 Hz data, is an accurate representation of the camera output operating at 300 Hz. In this case, integration prior to downsampling would not accurately reflect the operation of the particular sensor used in this experiment. Using this methodology, original 600 Hz data can easily be processed at 300, 200, 150, 120, 100, 75, and 60 Hz. Results of reducing the frame rate and computing the probability of detection are shown in Figure II.6. Here, the frame rates were reduced by removing frames. In each case, the starting frame was shifted in phase to represent all possibilities. Thus, if every other frame was used, generating 300 Hz data, the experiment was run 2 times with starting frames at 1, and 2. It is clear that high performance is maintained in the MWIR-1 and LWIR using frame rates as low as 100 Hz. The MWIR-2 reduces probability of detection more dramatically. One possible explanation for these reductions is that the algorithm is based upon assumptions of shape and relative positions of those shape features. Due to the fast rise times of hostile fire events, and depending on the phase shift of the starting frame, a removal of 6 frames could cause a notable change in the signature. The frame and time at which the peak occurs in the original sampled data may not be represented in the downsampled data. Depending on the phase, the peak may appear earlier or later, with respect to the full signature. Lower frame rate data also received fewer samples in the background population statistical window. Data with low SNR and exhibiting a noisy background, such as the MWIR-2 data, could have reduced accuracy in determining start and

stop times, affecting measured rise times and relative peak signal locations. Furthermore, frame rates at 75 Hz or lower, may not generate enough samples across the signature. The algorithm normalizes the signature to 10 intervals. Thus, signals with fewer than 10 samples are upsampled and as a result, shorter duration signals are more easily rejected. The algorithm was not optimized for performance at each frame rate. Performance can be improved by relaxing restrictions to accommodate the increased variability, but the increased probability of detection will also yield an increase in false alarms.

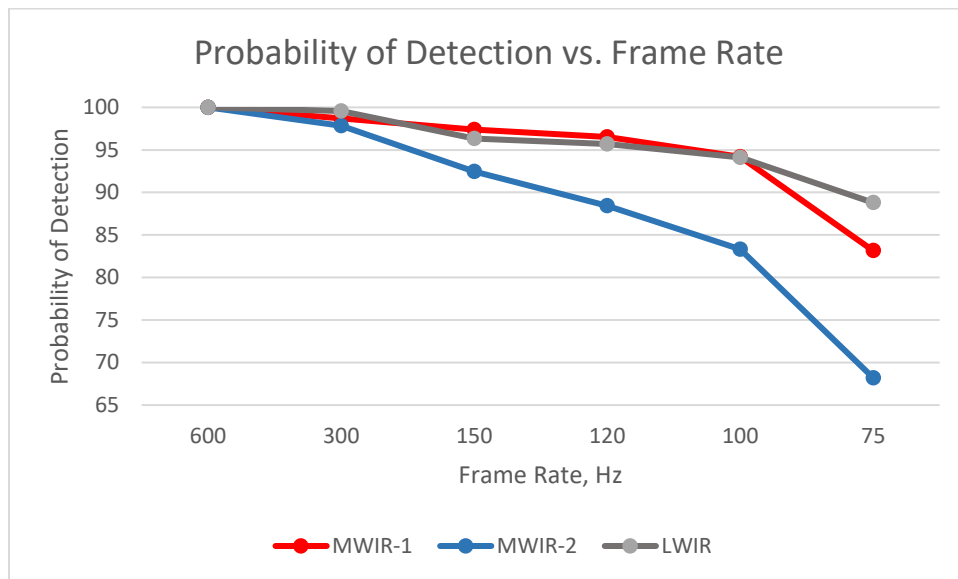


Figure II.6. Probability of detection plotted against frame rates for MWIR-1, MWIR-2, and LWIR broadband hostile fire data.

To better understand the effect of reducing the frame rate, false alarm data was recalculated using the same frame rate reduction process. Two false alarm videos, which were known to have a high number of false alarms in all bands were selected. Figure II.7 actually displays a reduction in false alarms at 300 Hz, which then begins to increase as frame rates are reduced further. These results were unexpected, but were helpful in identifying limitations on the implemented algorithm. As stated earlier, lower signal data near the noise floor may be

temporally clipped, yielding temporal shapes that do not resemble hostile fire signatures. Moreover, as frame rates are continued to be reduced, the number of frames required for a pixel value to be consistently above the background are reduced. Thus, the number of pre-detections drastically increases at lower frame rates.

To remove the effect of low signal noise and clutter, hostile fire data was empirically analyzed to determine average digital number (DN) counts above background. Outliers in the hostile fire data with reduced DN counts, such as out-of-focus data, were removed. Pre-detections which did not achieve one-third of the average hostile fire signature, were rejected. As a result, the unexpected decline in false alarms at 300 Hz is removed, as shown in Figure II.8. Though only the MWIR-1 data is shown, the MWIR-2 and LWIR data experience similar trends. At 75 Hz, it is theorized that the system does not receive enough samples to accurately detect targets, based on the detection parameter settings used. Note that the detection rate and false alarm rate are significantly reduced at this frame rate.

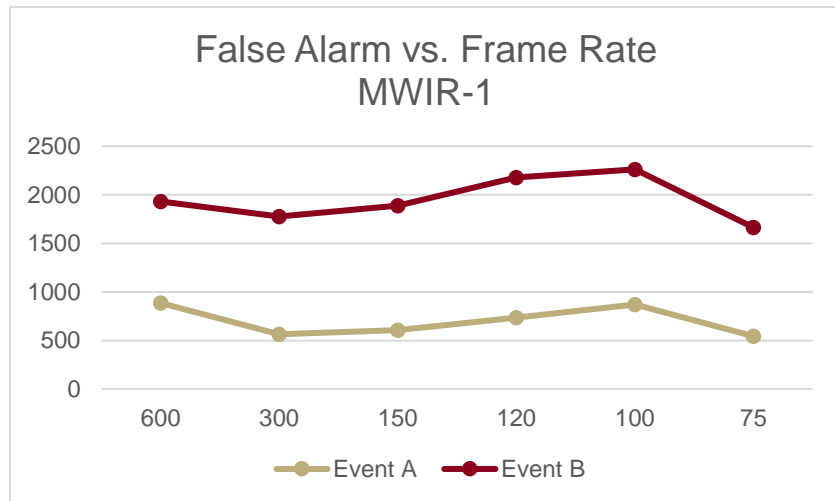


Figure II.7. Detected false alarms for varying frame rates. Low signal noise and clutter, and also reduced samples for background statistics and detection affect the false alarm rates.

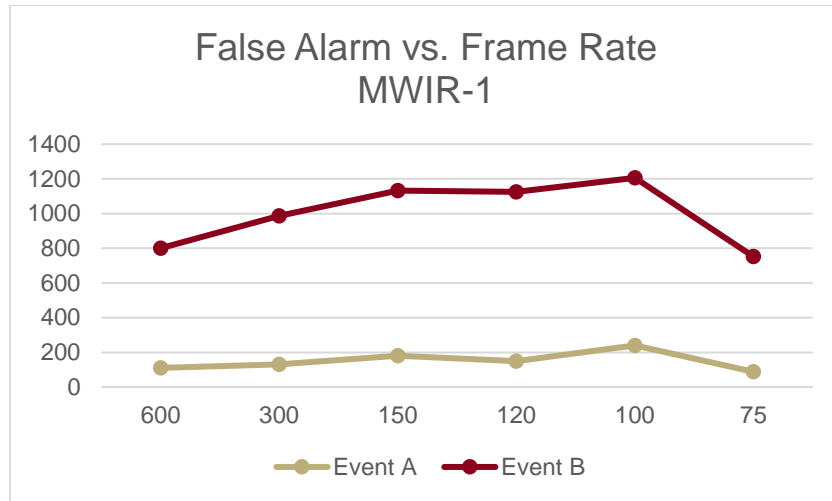


Figure II.8. Detected false alarms for varying frame rates. Low signal noise and clutter are removed.

To further reduce false alarms, multiple pixels in the neighborhood of the pre-detection pixel are analyzed. Pixels from a 3x3 neighborhood were summed at each frame to produce a neighborhood signature summation profile. In the same fashion, neighborhood pixels were ordered, and the median was extracted. Initial results of median profiles were poor and an average of pixels near the median was used to generate a temporal profile, instead. In terms of false alarm reduction, the neighborhood summation and median-average filters performed nearly identically. The neighborhood summation filter maintained 100% probability of detection in all bands, while the median-average filter had slightly reduced probability of detection rates of 99%, 98%, and 96% for the MWIR-1, MWIR-2, and LWIR, respectively. In considering the reduced probability of detection rate, increased processing, and negligible advantage in false alarm reduction, the median-average filter was not further pursued in this analysis.

Classification

Classification was performed using linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machine (SVM), and classification tree (Tree) classifiers. The models were evaluated using leave-one-out cross-validation. The classifiers

LDA, QDA, and Tree were trained using the 13 features, denoted as the 10% width, the 50% width, fall time, and the 10 temporally/intensity normalized samples, to discriminate between CLASS 1, CLASS 2, and CLASS 3. The results are shown in Table II.4. A layered approach was used for the SVM classifier, which first classified targets as either CLASS 1 or CLASS 2/CLASS 3. For this first layer, it was found that the duration features of 10% width, 50% width, and fall time were sufficient to discriminate CLASS 1 threats from other threats in the data set. A second SVM classifier was used to then differentiate the remaining threats of CLASS 2 and CLASS 3, using all 13 features. For comparison, the classifiers LDA, QDA, and Tree were also used in a binary layered approach, mimicking the features used for the SVM layering. These results are also provided in Table II.4. For the total probability of classification, prior class distributions were taken into account as weighting factors for each layer. The total probability of classification can then be compared to the single layer probability of classification, but the results of each layer can help to identify potential difficulties in individual discriminations. For this data set, QDA and SVM showed the greatest level of performance amongst the classifiers evaluated. From the results, it can be observed that CLASS 1s are more easily classified for this data set, while CLASS 2 and CLASS 3s are less separable.

Table II.4. Classification results of LDA, QDA, Tree, and SVM using leave-one-out cross validation.

Classification Results						
	<i>Threat Classes</i>	<i>Features</i>	<i>LDA</i>	<i>QDA</i>	<i>Tree</i>	<i>SVM</i>
Single Layer	Class1,Class2, Class3	Duration and Shape	93.97%	95.69%	92.24%	N/A
Double Layer (1)	Class1, Class2&3	Duration	92.24%	99.14%	100.00%	100.00%
Double Layer (2)	Class2,Class3	Duration and Shape	96.97%	98.48%	86.36%	93.94%
Total (1 & 2)			93.96%	98.90%	95.05%	97.80%

The results of the preliminary classification analysis are promising, but assume that all inputs are one of the three hostile fire classes. In order to further reject false alarms from the detection stage, an additional false alarm class can be established for classification. The space of potential false alarms is unknown and can be difficult to define. False alarms in one environment may display different characteristics than false alarms in another environment. Two approaches to classifying the unknown false alarm class are presented here. One approach, defined here as “dynamic distance,” attempts to adapt the false alarm class to the changing environment, while the second approach, defined here as “layered distance,” relies on the hostile fire class being accurately represented.

Providing a representation of possible sources of false alarms in varying prospective environments can be a challenge due to incomplete knowledge of all possible false alarm sources. The concept of the first approach is to provide a representation of the false alarms in the current environment and adapt to changing environments. In this first approach, it is assumed that the system is provided with a set of initial false alarms from the current environment. This false alarm seed could be determined during an initialization period. The initial false alarm seed is stored in a buffer. As new inputs are classified as false alarms, they are stored in the buffer, replacing older false alarm observations. The buffer, therefore, stores the most current representation of the false alarm class. The classifier is then reconstructed periodically to reflect the current false alarm class and the static hostile fire classes (see Figure II.9). This dynamic false alarm adaption assumes that the current scene false alarms can be represented by a single class. A possible improvement, such as using a k-means clustering, could entail dynamically adapting the number of false alarm classes to provide the greatest clustering of false alarm

classes, while minimizing the number of false alarm classes. There is also the assumption that the false alarm class will not accept and incorporate real hostile fire into the false alarm class.

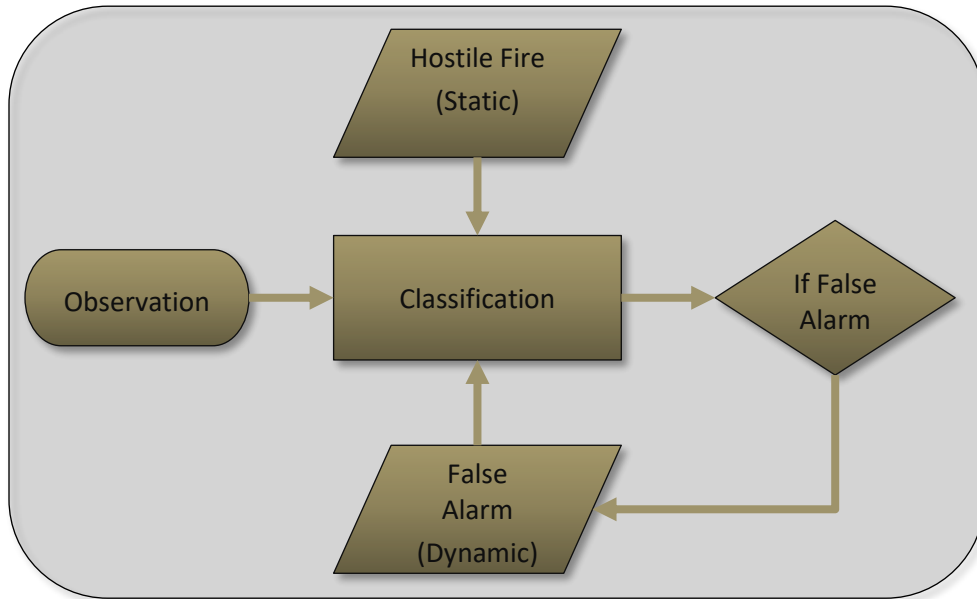


Figure II.9. Classification scheme “dynamic rejection” incorporating dynamic false alarm class updating.

The second approach does not rely on knowledge of the false alarm class, but rather relies on knowledge of the hostile fire class. For the results in Table II.4, the solution space is partitioned into three regions, reflecting the hostile fire classes. Thus inputs, which may not be similar to any of the three hostile fire classes are forced into one of those three known classes. The second approach applies a distance metric to the input in conjunction with the layered classification scheme shown in Table II.4. The goal is to reject inputs that are not similar to any hostile fire class, while still discriminating between real hostile fire events. A binary layered decision tree, similar to the layered approach discussed previously, is used for classification. The approach is modified so that at each classification of threat, a distance value is measured from the threat class. Inputs classified as a particular threat, but which are dissimilar from the threat, are rejected and labeled as false alarms. The decision scheme is illustrated in Figure II.10. The

distance metric, the normalized Euclidian distance, uses all 13 features previously discussed, and is defined as

$$\text{Normalized Euclidian Distance} = \sqrt{\frac{\sum_{i=1}^n \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2}{n}} \quad (2.3)$$

where x_i is the i^{th} of n observation features. Additionally, μ_i and σ_i are the mean and standard deviation of the corresponding population feature. The classification approaches were tested against the events A and B, used in the frame rate reduction analysis. The QDA classifier was chosen due to high performance results in the three class problem, but is not inherent in either approach. In Table II.5, the results of the classification approaches are provided. The initial false alarms generated from the detection stage, using the summation filter, are also listed for comparison. An improvement in false alarm reduction is evident by the results. In testing against the actual hostile fire data, a correct classification rate of 99.08% was achieved by the dynamic rejection approach, and a correct classification rate of 98.28% was attained by the layered distance approach.

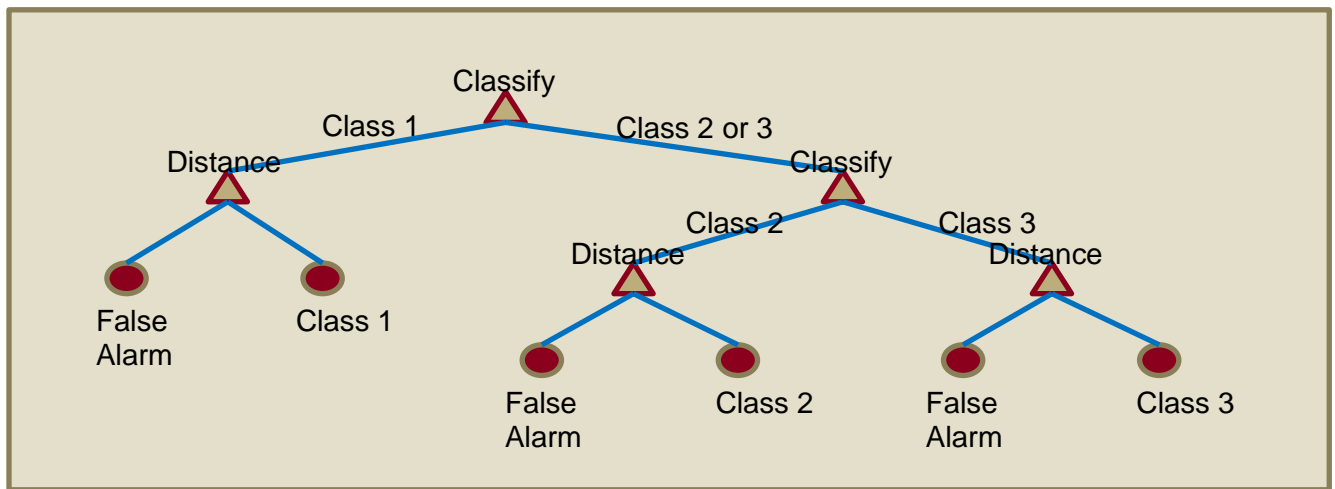


Figure II.10. Classification scheme “layered distance” to include false alarms as an additional class by measuring the distance from the respective class.

Table II.5. False alarm reduction from the detection stage, using an adaptive false alarm class approach, and a distance from hostile fire approach.

LWIR 600 Hz False Alarms			
	<i>Detection Stage</i>	<i>Dynamic Rejection</i>	<i>Layered Distance</i>
<i>Event A</i>	217	3	13
<i>Event B</i>	3623	356	220

Discussion

Ground, on-the-move, hostile fire detection, with near and far field scenery is a complex problem. Real hostile fire data was collected in multiple environments, which included various scenery, temperatures, and humidity. For each threat class, at least two types are represented in the data set and each threat class appears in at least two environments. On-the-move false alarm data was also collected. An algorithm using low computational complexity temporal and shape features, such as duration, rise time, and integrated samples, was shown to have high probability of detection, but also high false alarm rates. MWIR bands yielded significantly fewer false alarms, but showed a dramatic increase due to sun glints. The LWIR band provided better imaging of terrestrial scenery, but also suffered false alarms due to motion. Features measured from LWIR hostile fire demonstrated consistency and were suitable for performing classification. Threat classification for the three-class problem of distinguishing hostile fire, was extended to include a false alarms class. Two methods were provided for potentially removing false alarms at the classification stage.

An analysis was conducted to determine the suitability of reducing frame rates. Without optimization the algorithm was tested at lower frame rates. Due to the longer time constant of the cameras used, hostile fire signatures are observed for longer durations. Even at reduced frame rates as low as 100 Hz, hostile fire signatures could be detected at rates greater than 94% for

MWIR-1 and LWIR. An effect of the time constant is also reduced signature amplitudes. MWIR-1 and LWIR yielded higher signature amplitudes than the MWIR-2. The algorithm was found to be sensitive to the reduced signatures of the MWIR-2 band and, as a result, performed poorly in this band at reduced frame rates. False alarms were reduced after filtering low signal noise and clutter, but this filtering process makes an assumption about the system's anticipated DN counts for hostile fire. Additionally, neighboring pixels were considered for evaluation. At longer ranges, hostile fire will become sub-pixel, and detection of signatures in neighboring pixels is a result of optical blur and signal intensity. These filtering processes were successful in removing false alarms, but could impact maximum range performance, as they rely on signal strength.

Though the false alarm problem is not completely solved using these techniques alone, there are a few observations which can be made. Lower frame rate cameras are more commercially available. The high frame rate data for this work was collected at the cost of windowing the field of view. A larger field of view can provide greater situational awareness. The LWIR band, used for imaging terrestrial scenes and providing situational awareness, also displayed an ability for capturing high intensity, consistent signatures suitable for detection and classification. The LWIR severely suffered from false alarms when used for on-the-move hostile fire detection. The MWIR-1 band generated far fewer false alarms, presenting itself as a better candidate for detection. A two-channel system, using MWIR-1 for detection and LWIR for classification and situational awareness could be feasible. LWIR data could be used in further discriminating false alarms, while MWIR data features could help to improve classification. Additionally, registration could potentially be performed on LWIR data improving false alarm reduction. These are areas that will be under further study in the future.

III. CLUTTER MITIGATION METHODOLOGIES

Introduction

In this chapter, clutter rejection techniques are addressed. The detection process, identified in Chapter 1 is adopted. Beyond the detection stage, a classifier scheme is presented, which not only classifies weapon fire by weapon type, but also mitigates clutter in a dynamic environment. Methods are also provided to minimize processing and memory requirements, making the methodology practical in application.

Problem description

An improved methodology for practical implementation of weapon fire detection. Demonstrates an improved methodology for increased clutter rejection, making said approach feasible. Also provides methods to allow for real time implementation, using practical sensor parameters (frame rate, sensitivity, processing computations).

Methods

Signatures

In this paper, a hostile fire detection and location algorithm is described in which temporal duration and shape are exploited. Hostile fire temporal signatures analyzed exhibit a fast, sharp, rise in intensity, followed by a slower, exponential decay. The exponential decay is partially due to the time constant of the sensors used to collect the signatures for this study. These signatures were collected using a LWIR, MWIR-1, and MWIR-2 sensor operating at 600 Hz. An example of temporal signatures is shown in Figure III.1. The signatures were collected at 600 Hz frame rate, and downsampled to 120 Hz. Hostile fire signatures can demonstrate significant variability for different threats, which can increase the complexity of target discrimination. Examples of this variability include saturation, signal from booster motors,

elongated tails, and noisy signals. At 120 Hz, the high frequency effects of noise and saturation are reduced, however, information and signal fidelity are also reduced. A successful hostile fire algorithm must handle these variabilities.

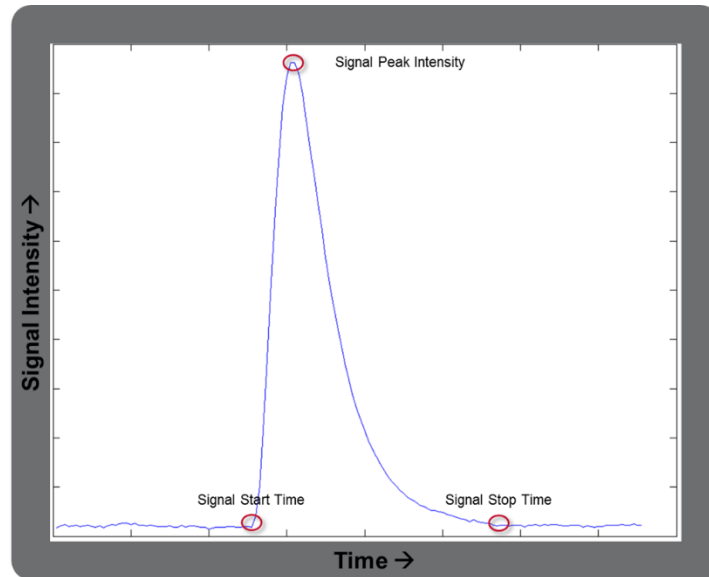


Figure III.1. An example of hostile fire signatures.

Algorithm

Introduction

The hostile fire detection and location algorithm presented in this paper builds upon the work of Chapter 2, in which an initial detection scheme using single band imagery is provided along with potential clutter mitigation techniques using classification schemes. This chapter will focus on work which extends the detection and classification process, and the reader is referred to the prior chapter for a detailed explanation of these previously described methodologies. The algorithm presented in this chapter is composed of three stages, pre-detection, detection, and classification.

Pre-Detection

The pre-detection stage identifies pixels which have consistent signal above the background. In determining signal, a background model is developed, and a signal-to-background ratio (SNR_B) is computed. For each pixel, a temporally sliding window is used to establish mean (μ_{bkg}) and standard deviation (σ_{bkg}) statistics of the background. The sliding statistical background window is separated from the current frame by a temporal buffer. The separation provided by the buffer allows for both fast and slow changing signals to be sensed for elevated SNR_B , where SNR_B is defined as

$$\text{SNR}_B = (\text{pixel}_{\text{current}} - \mu_{bkg}) / \sigma_{bkg} \quad (3.1)$$

To more easily identify hostile fire across multiple pixels, when apparent scene motion is present, each incoming frame is first preprocessed using a local maximum filter. Pixels which vary from the background are tracked to determine whether a consistent signal persists over a duration of time. Pixels which exhibit persistent signal above the background are recorded as pre-tracks. These recorded pre-tracks are evaluated after a set period of time (m -frames), in which any potential hostile fire signature would complete. Thus, at each frame, pre-tracks can be recorded and each pixel has the potential to be a pre-track during these m -frames. To minimize the number of recorded pre-tracks, the pre-tracks of each frame can be associated or grouped with neighboring pre-tracks. This grouping can be performed both spatially and temporally, reducing the number of recorded pre-tracks to be evaluated.

An alternate option to grouping, is to use an n -window memory system, where each window is the size of the image frame. Each window is able to record the pixel location of pre-tracks and when they occurred. For an n -window system, n pre-tracks can be recorded at each pixel location. For $n < m$, there is the possibility that not all pre-tracks will be recorded. The aim

is to reduce the recorded pre- tracks to a limited, fixed number, yet maintain pre-tracks from actual hostile fire events. Given that the occurrence of real hostile fire signatures dominate the statistical background and suppress new pre-detections from occurring on a pixel during the event, and using knowledge of hostile fire event durations, a minimum duration threshold between pre-tracks can be determined. Thus, if two pre-tracks occur on the same pixel within the threshold period, the first occurrence of the two pre-detections can be discarded immediately. This process, in effect, filters out high frequency noise and clutter. The n -window system is similar to an ordered list. The most recent recorded entry can be overwritten if it is deemed noise, and the n^{th} entry is repeatedly overwritten while the ordered list is at capacity (see Figure III.2). Entries are normally removed when they are analyzed m -frames after being added. To ensure that hostile fire events are not overwritten, the number of required windows, n , can be estimated by

$$n \geq m / \text{minimum duration threshold in frames} \quad (3.2)$$

The minimum duration threshold takes into account the (minimum) expected duration of hostile fire events, as well as the statistical background window, buffer window, and the pre-track persistence threshold. If expected durations of hostile fire signatures vary significantly, the process should accommodate longer duration events in the m -frame evaluation period, and the potential of multiple smaller duration events occurring in the same m -frame evaluation duration. Alternately, the evaluation can occur in parallel, such that an n -window memory system is used to evaluate longer duration events within an m -frame evaluation period, and an alternate n -window memory system is used to evaluate shorter duration events within an m' -frame evaluation period. This can allow shorter duration events to be processed with less latency. The window system sets the maximum memory storage which is used to store pre-detections. In

Chapter 2, it was observed that a significant number of pre-tracks can occur. A non-prioritized and unbounded memory assignment of pre-tracks can lead to reduced processing time, and also lead to missed detections.

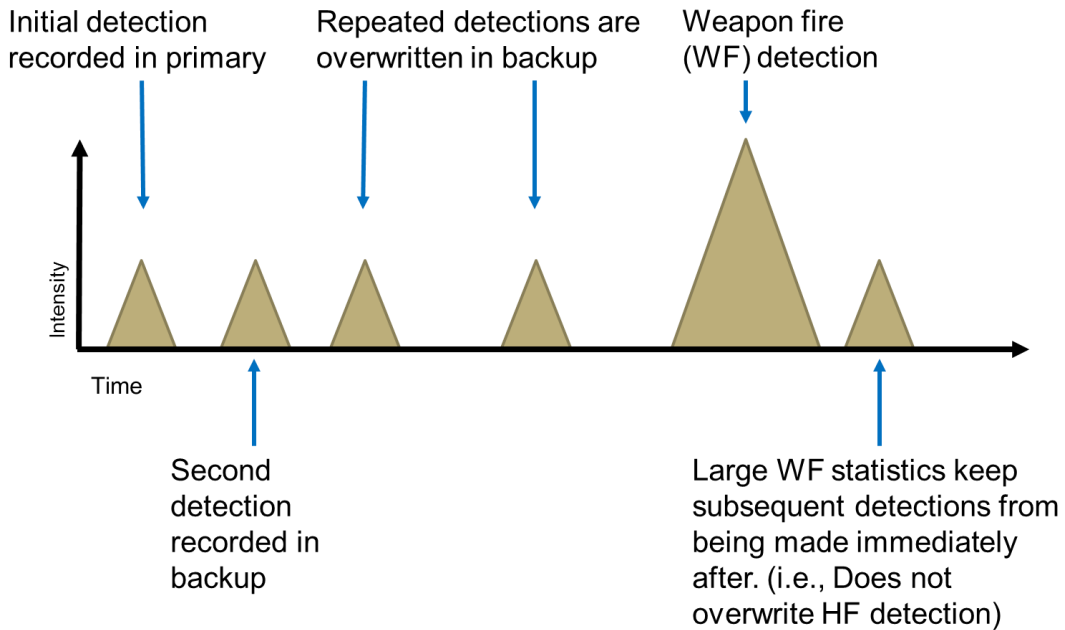


Figure III.2. Diagram depicting pre-detection memory storage process using an n -window memory system, where $n = 2$, (1: primary and 2: backup). Here, the second detection is greater than the minimum duration threshold. Subsequent detections are overwritten in the backup window.

Pre-tracks are evaluated after a set period of time, in which the hostile fire signature is allowed to complete (see Figure III.3). The evaluation is as described in Chapter 2. It should be noted that concept of selecting a maximum pre-track within a neighborhood is analogous to grouping multiple pixels, with the intent of returning a single representative pixel. This serves the dual purpose of accurately representing targets and reducing false alarms, as well as reducing the amount of data for further processing. The condensed data reduces memory usage and processing times.

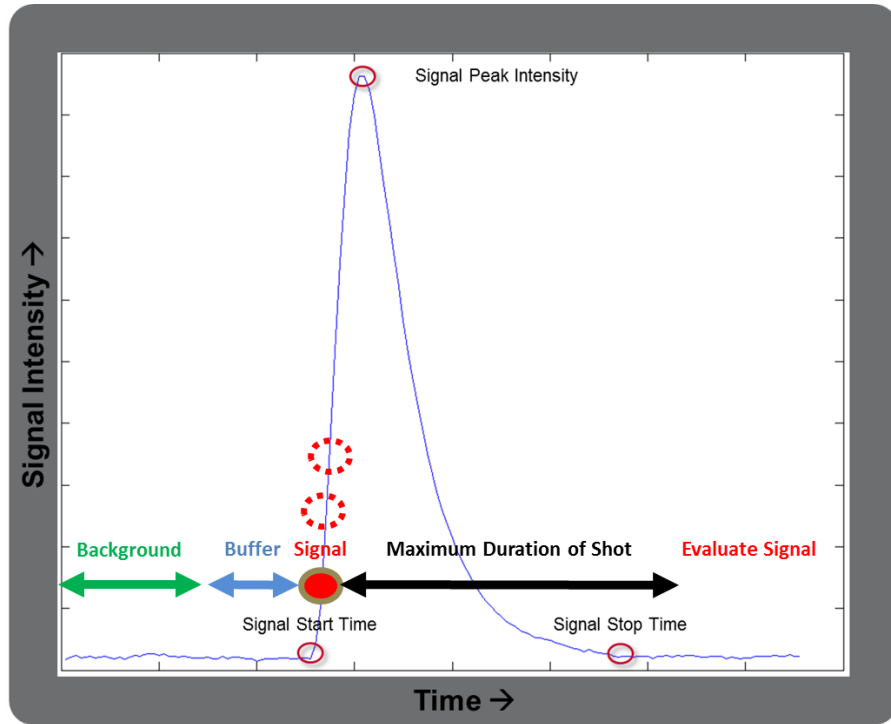


Figure III.3. A depiction of the pre-detection evaluation process, which assesses the pixels that deviate from the statistical background after a period of time.

Detection

At the detection stage, potential hostile fire signatures are extracted from pre-detections. Pre-detection pixels are investigated within a temporal window surrounding the peak signal intensity value. The temporal window is determined using knowledge of hostile fire signature durations. The signal within the temporal window is extracted. The extracted signal is measured to determine the actual start and stop times of any hostile fire signature within the window. Accurate measurements of the start and stop times ensure that the signature is extracted completely without additional unwanted background clutter. The statistical background and buffer system is applied to determine start time. Stop time can also be found using the same method, where background is established post-shot and the signal is retraced in reverse. The start time is verified by investigating changes in slope near the initial rise of the hostile fire signature.

Additionally, measuring when the signal rises and returns to a percentage of the peak value, is used to verify the consistency of start and stop times. Though noise can often be found in the tail of hostile fire signatures as the signal exponentially decays toward the background, an averaging filter can be employed to investigate slope and intensity changes to verify the stop time. Accurate, complete, and consistent hostile fire signature extraction forms the basis for signature metrics and features.

An extracted temporal signature, within measured start and stop times, is analyzed for duration and shape features. The signature rise time to the peak and the fall time of the exponentially decaying tail are measured. The widths at 10% and 50% of the signature peak value also yield measures of duration. To measure shape, the signature is temporally resampled into a standard number of samples and then normalized in intensity, as shown in Figure III.4. It was determined that a 10 sample total was sufficient to mitigate noise yet maintain general characteristics of shape. The temporally resampled and normalized signatures provide reliable measures, such as the greatest change in consecutive samples, the distribution of energy, presence of a monotonically decreasing tail, and an exponential decay within a portion of the tail. The change in intensity between the background and signature peak is also measured, to ensure that low intensity noise and clutter are discarded. Thus duration, temporal shape, and intensity scale are captured. The methods described in this dissertation primarily address temporal aspects of hostile fire signatures. A measure of the spatial characteristics of hostile fire can be found by summing spatially neighboring pixels of the extracted signature. The summation is performed at each frame, generating a temporal summation profile. The summation profiles of real hostile fire events exhibit similar duration and shape characteristics to single pixel profiles. To reduce processing of unwanted noise, the signature's background is removed, peak intensity value is

measured and low intensity signals are removed. Signatures which meet threshold criteria for the detection measurements described are considered detections.

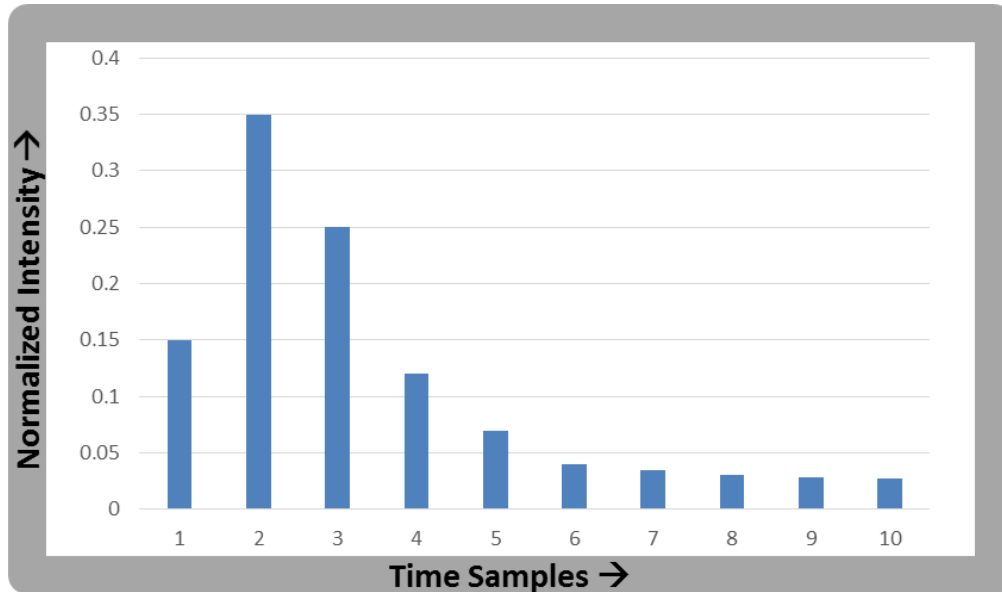


Figure III.4. An illustrative example of a temporally and intensity normalized signature.

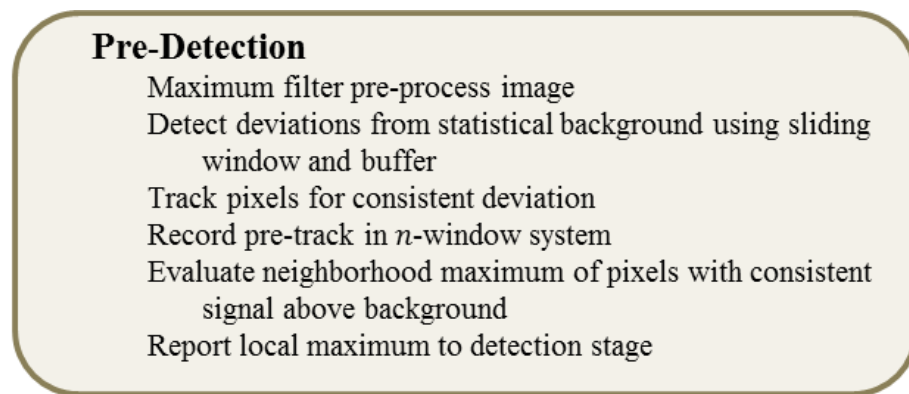
For practical usage, a single hostile fire event, should generate a single hostile fire declaration. For closer hostile fire events spanning multiple pixels, multiple detections can occur. Recall that pre-detections were spatially grouped within a local neighborhood at each frame. Employing the method as described previously allows for pre-detections to be generated at each frame. A method analogous to temporal grouping is described here, which reduces redundant information and processing by analyzing detections temporally to determine the best representative candidate. In real hostile fire data, though, multiple pixels may exhibit characteristics of hostile fire signatures. These characteristics are often less reliable towards the outer spatial extent of the signature. The best candidate detection is chosen by maximum peak intensity value. Alternate measures can be used, such as the change in intensity, or even voting schemes using combinations of intensity and duration measurements. Processing performance, reliability, and hostile fire phenomenology led to the choice of maximum peak intensity as the

criterion, however, alternative measures require further research to determine their effectiveness and suitability. This method is not suitable for significantly saturated signatures, where saturation occurs spatially and temporally.

The method described thus far defines a detection scheme for a single band. Thus, for each band in a multi-band system, pre-tracks, pre-detections, and detections are made. Amongst these detections, the best representative detection candidate is chosen. For dual band or multi-band application, a single detection should be returned, given a single hostile fire event. Hostile fire detections across bands or sensors can be associated using time of occurrence of the hostile fire detection, such that a detection generated in one sensor will prompt a search for detections in other sensors for the same period. One approach is to use a guard band. Only after detections have occurred in guard band sensors are other sensor streams searched and evaluated. This procedure can significantly reduce processing. For example, the LWIR imaging sensor can generate a significant number of pre-detections and detections, due to scene motion when on the move. Because the MWIR-1 sensor produces low intensity features of targets with temperatures around 300° K, such as typical terrestrial scenes, fewer pre-detections and detections are generated as a result of scene motion, making the MWIR-1 favorable for use as a guard band. Consequently, a detection in the MWIR-1 band can be used as an indicator to search the LWIR band for potential detections at the corresponding time and location.

In the data collected for algorithm development, the sensors were not frame synchronized, and therefore a temporal search window was utilized. Though frame synchronization will allow for narrower search windows in practice, the process may still account for variations in signatures between bands. Additionally, for sensors with overlapping fields of view, a spatial search window can also be implemented. The sensors which captured the

data used for algorithm development were not optically registered to tight tolerances. Due to the low intensity terrestrial scene images generated by the MWIR-1 band, spatial registration is not feasible. Therefore, a spatial search window was applied. Here, improved optical registration will yield a narrower search window. A spatial search window is still recommended to accommodate for imperfections in optical alignment, aberrations, and variations in muzzle flash signatures of hostile fire events between sensors. Detections determined in the MWIR-1 guard band and the LWIR verification band, which have been associated in time and space, are considered to be dual band detections, and passed to the classification stage for further analysis. Quick reference guides for pre-detection and detection are provided in Figure III.5 and Figure III.6, respectively.



Pre-Detection

- Maximum filter pre-process image
- Detect deviations from statistical background using sliding window and buffer
- Track pixels for consistent deviation
- Record pre-track in n -window system
- Evaluate neighborhood maximum of pixels with consistent signal above background
- Report local maximum to detection stage

Figure III.5. Quick reference guide for the pre-detection scheme used.

Single Band Detection

- Extract shot signature (start / stop time)
 - Deviation from background, change in slope at rise, signal return to percentage of peak, slope of averaged tail
- Measure signature duration
 - Width at 10%, 50% of peak
- Measure shape
 - Temporally / intensity normalize signature into 10 samples
 - Peak location, decreasing tail, energy distribution, rise slope, decreasing Δ slope
- Measure amplitude
 - Change in digital numbers (DN) from background
- Measure neighboring pixels
 - Neighborhood summation
- Select best in-band candidate detection within a spatial / temporal neighborhood
 - Maximum peak intensity

Multiple Band Detection

- For each guard band detection, search the verification bands for detections within a spatial / temporal neighborhood.

Figure III.6. Quick reference guide for the detection scheme used.

Classification

The classification stage serves a dual purpose by classifying hostile fire events into classes of threats and rejecting potential false detections. Chapter 2 demonstrated that hostile fire events could be classified at 600 Hz and provided two potential schemes for using classification to remove false detections. One scheme, labeled dynamic rejection, used a classifier which included a false alarm class as well as hostile fire classes. This scheme included a feedback system in which detections, declared as false alarms by the classifier, were fed back into a dynamically updating false alarm class. The classifier periodically updated to reflect changes in the false alarm class, which implied an adaptation to changes in the environment generating the false alarms. The other scheme, termed layered distance, relied on binary classification using a binary tree. Thus CLASS 1s were classified against all other threats. The other threats were then classified as either CLASS 2 or CLASS 3. To remove potential false detections, a distance

measure was applied to the results to determine whether the detection, classified as hostile fire, was similar to the respective hostile fire class. Both schemes, showing promise in their ability to reduce false alarms, have been combined here in a two-stage classification scheme approach, which shall be referred to as dynamic distance.

The two stage dynamic distance classification scheme segments false alarms from hostile fire in the primary stage, and then classifies the resulting hostile fire into specific classes, after which a distance metric is applied to verify membership in the specific hostile fire class, visualized in Figure III.7. The first stage employs a multi-class classifier, which segments false alarms from hostile fire. False alarm detections update the false alarm class and the classifier itself is periodically updated. To ensure stability, the false alarm class can be partitioned so that a percentage of the population is updated dynamically, while the remainder is static and remains unchanged. Additionally, multiple false alarm classes can be used to represent all false alarms. In a preliminary investigation of false alarm data, false detections were groundtruthed according to the apparent scene false alarm source. The results indicated false alarm classes, such as a “sun glint” class, that would apply across various days, environmental conditions, and scene content, could not be made based on the data. Additional research is required on false alarm signatures and the complex processes by which they are generated. Unlike the false alarm class, the hostile fire classes are static, and generated from known hostile fire data. Though detected events can be classified as various types of hostile fire, the primary goal of the first stage is simply to remove the majority of false alarms. While false alarms are fed back into the false alarm class, detections classified as hostile fire in the primary stage are passed to the secondary classification stage.

The secondary classification stage accepts hostile fire detections from the first stage, and classifies these detections into their respective types of hostile fire. Here, only hostile fire classes

are used for classification. This process can be implemented in different ways, such as using a single classifier incorporating all hostile fire classes, or using a binary tree with multiple binary classifiers. Because the false alarm class of the first stage can potentially overlap with hostile fire classes in decision space, a more accurate decision boundary tends to occur between hostile fire classes when only hostile fire classes are used to generate the classifier.

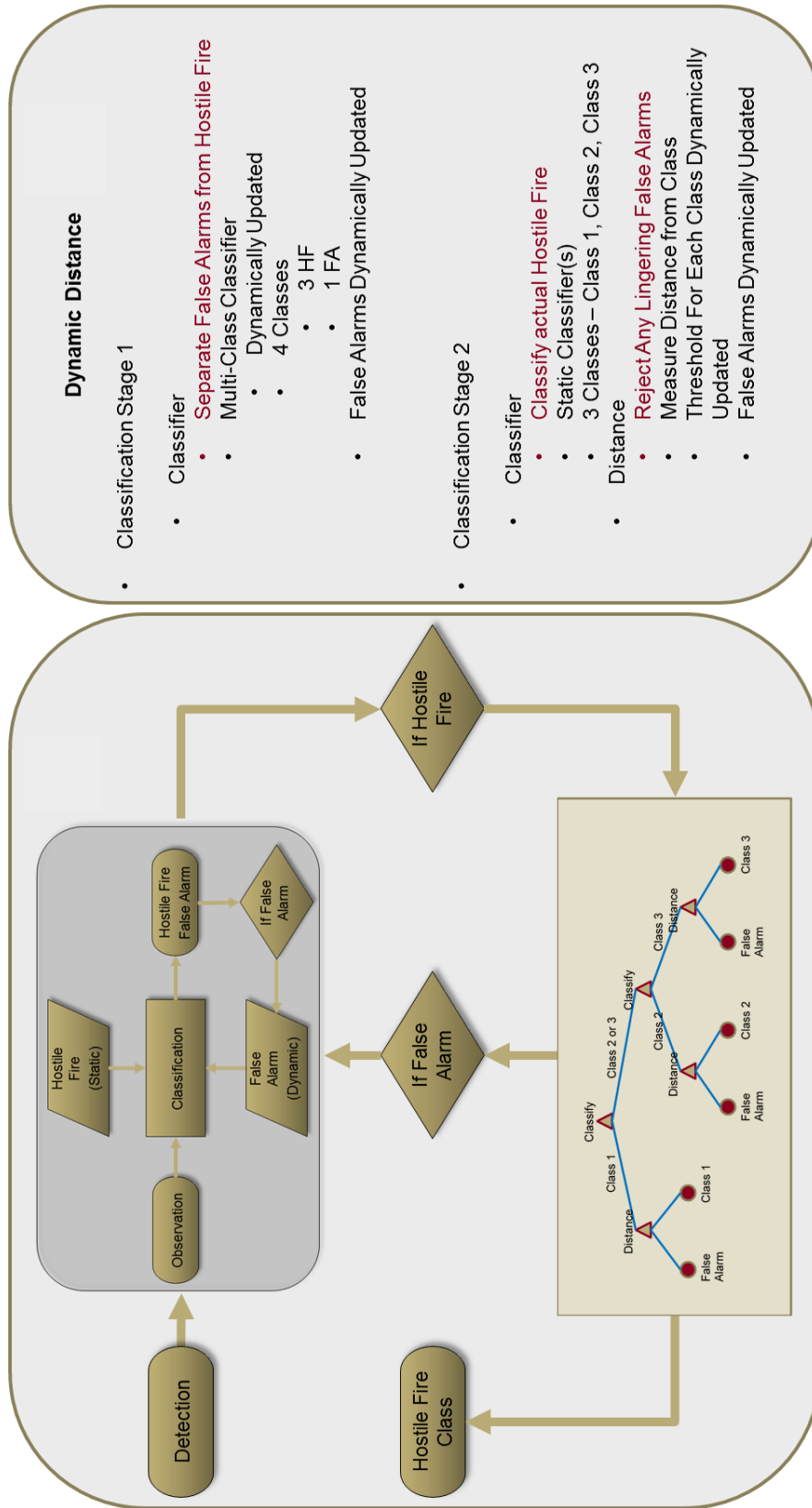


Figure III.7. Dynamic Distance classification scheme.

In practice, however, false alarm detections can propagate to the second stage of classification. In this second stage, classifiers are not used to reject false alarm detections, but instead classify all remaining detections as one of the hostile fire classes. In order to reject false alarms, a distance measure is applied to determine whether the current detection is similar enough to its newly classified hostile fire class. The class cluster distance, which can be interpreted as a normalized Euclidian distance, measures features of the detection, and generates a single value representing the distance from the center of its hostile fire class. The class cluster distance is given by

$$\text{Class Cluster Distance} = \sqrt{\frac{\sum_{i=1}^n \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2}{n}} \quad (3.3)$$

where x_i is the i^{th} feature of the observation, n are the number of features, and μ_i and σ_i are the class mean and standard deviation of the i^{th} feature, respectively.

The concept is that observations which belong in the class will have a smaller distance from the class centroid than observation which do not belong to the class. The class cluster distance reduces n features or dimensions of data to a single dimension or distance. Because all n features contribute to the distance measure, the features should be chosen which reduce the distance of hostile fire from their respective hostile fire class, yet extend the distance of false alarm detections from the hostile fire class. Therefore, features are chosen which provide the greatest separation between the hostile fire and false alarm classes, so that an incoming false alarm detection will be separated from the hostile fire classes. Here, there is an assumption that a false alarm detection is similar to the false alarm class. The separation distance between two classes along a given feature is provided by

$$Feature\ Distance = \frac{abs(\mu_1 - \mu_2)}{\sqrt{1/2 (\sigma_1^2 + \sigma_2^2)}} \quad (3.4)$$

where μ_1 and σ_i are the mean and standard deviation of the feature in question for class i , and $i = 1,2$. Here, the two intended classes are the hostile fire and the false alarm classes. Because the features are chosen by their ability to separate the hostile fire and the false alarm classes, and because the false alarm class is dynamically changing, the chosen features are periodically updated as well. In detection theory, the sensitivity index or d' , provides a similar measure of separation.

The distance measured by the class cluster distance is used to determine whether observations are considered hostile fire or rejected as false alarms. A threshold value is used to provide this discrimination and can be chosen using various methods. In theory, well-chosen features will yield a tightly clustered population. In practice, outliers may exist in the population. Extending the threshold to include these outliers may reduce the capability of discriminating false alarm detections. One potential method of choosing a distance threshold is by selecting a threshold which accepts a percentile of the known population. For hostile fire detection, it is undesirable to miss actual hostile fire events, and a minimum threshold can be chosen which accepts 100% of known hostile fire events in the population. This threshold can be further scaled by $\alpha > 1$, which implies that the known population is not robust or representative of all hostile fire of a particular class. Scaling the threshold by a value greater than 1 allows for variations in signatures which are not represented in the data set. For the distance measure to be effective, it is assumed that actual hostile fire detections are classified into the correct class. For example, a CLASS 1 observation classified as a CLASS 3 observation will be compared in distance to the CLASS 3 class and not the CLASS 1 class. For this example, if significant difference exists

between the CLASS 1 observation and the CLASS 3 class, a well-suited classifier will not classify the CLASS 1 observation as a CLASS 3. Figure III.8 displays the distance threshold and feature determination process.

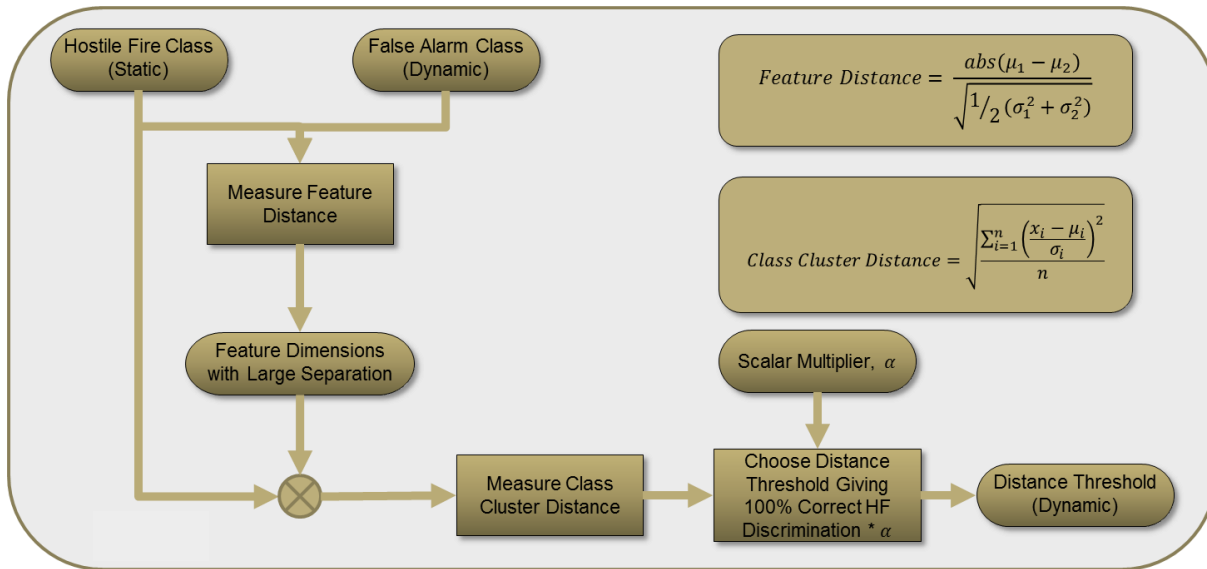


Figure III.8. The distance threshold is determined dynamically using the static hostile fire classes and the dynamic false alarm class.

Results

Detection

The hostile fire detection algorithm was tested against real imagery of threats consisting of CLASS1, CLASS 2, and CLASS 3 hostile fire. Data was collected for these threats over the course of six data collections. The environments consisted of freezing temperature and snow, warm desert, and moderate green fields and tree lines with varying humidity which extended from dry desert to daily thunderstorms. Each threat class appeared across at least two of the three environments. Within each threat class, there existed at least two variants of each threat type. In total, 122 hostile fire events were used in this analysis. Additionally, data was collected of various scenes without threat targets, both stationary and on-the-move, which was used to perform false alarm analysis. The false alarm data collection was conducted over the course of

three days, the first two of which were cloudy and the last day was sunny. The scenarios captured on the first cloudy day, were repeated on the sunny day for comparison. The false alarm scenarios included sun glints from water, sun glints from cars, headlights from passing cars, sodium lamps, the sun passing behind tress, welding, and a burning barrel with flares.

The data was collected using imagers consisting of two sub-band MWIR imagers, MWIR-1 and MWIR-2, and an LWIR. These 17 μm , 640 x 480 pixel imagers were run at 600 Hz, with a 40° FOV lens. To attain the desired high speed frame rate, the imagers were windowed to 640 x 44 pixels. Though the original data was recorded at 600 Hz, the data was analyzed at various frame rates from 600 Hz to 100 Hz. To achieve lower frame rates, the data was downsampled, such that 120 Hz data was generated by removing 4 out of every 5 frames. In this case, however, the choice of starting offset can yield 5 unique representations of the data. Figure III.9 provides a graphic depicting this example. To ensure that an artificial bias was not created, each potential offset was analyzed for data containing hostile fire events.

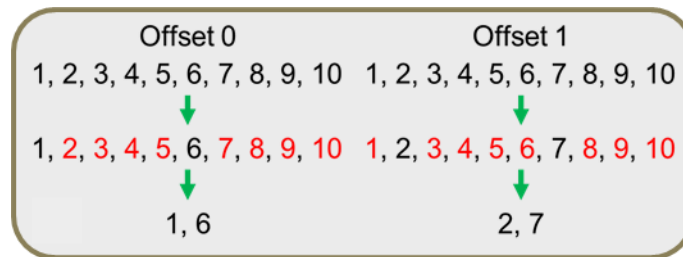


Figure III.9. An example of down-sampling and the choice of offset.

An analysis was performed on detection rates as a function of frame rate. The analysis was conducted using the LWIR, MWIR-1, and MWIR-2 bands, as well as a LWIR/MWIR-1 dual band combination. The detection rates reflect the detection stage only, and do not incorporate classification results here. In Figure III.10, it can be observed that overall detections rates decline as frame rates are reduced. The LWIR and MWIR-1 bands perform well for frame

rates as low as 100 Hz. The dual band LWIR/MWIR-1 requires a detection in both individual bands, thus the performance can never be higher than either of the individual bands. The requirement of dual band detection does not significantly reduce detections rates for LWIR and MWIR-1, and yields a detection rate of 97.7% at 120 Hz, and a 100% at 600 Hz. The MWIR-2 performs satisfactorily, but the impact of lower SNR is observed in the reduced detection rates. At frame rates lower than 100 Hz, the number of temporal samples across hostile fire signatures do not meet algorithm parameter requirements. In particular, the choice of the number of temporally integrated and normalized samples significantly affects these results. Detection rates, however, do not provide a complete view of performance, and must be considered alongside classification rates and false alarm rates.

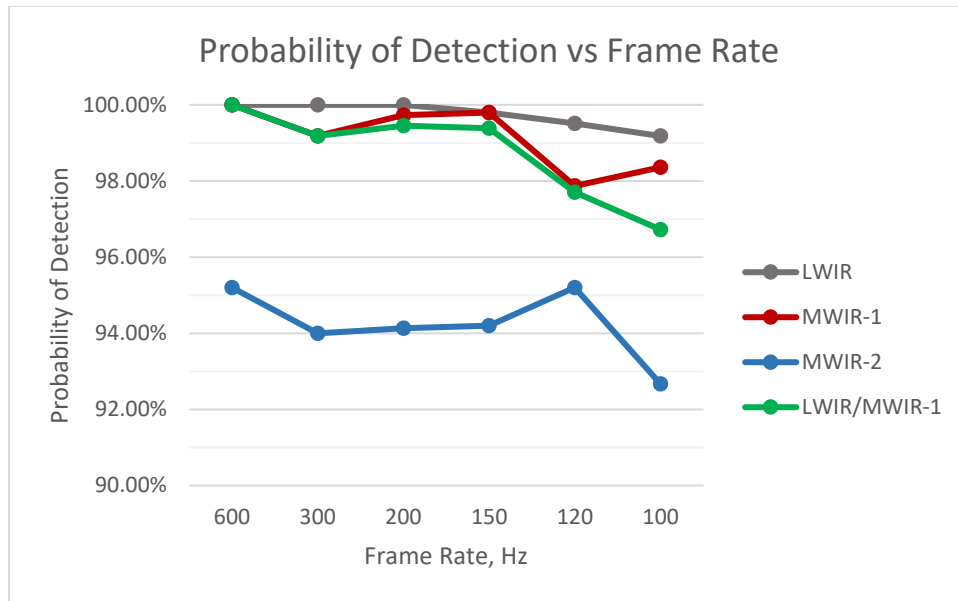


Figure III.10. The probability of detection as a function of frame rate for LWIR, MWIR-1, MWIR-2, and LWIR/MWIR-1.

A similar analysis was performed on the false alarm data set. A subset of the false alarm data set was processed for detections using the LWIR, MWIR-1, and MWIR-2 band, and also the LWIR/MWIR-1 dual band. This subset of data uses the sunny day of data which was shown to

be the worst-case scenario and generated more false detections for analysis. In Figure III.11 the LWIR and MWIR-1 false detections are displayed as a function of frame rate. The results have not been scaled to reflect false detections per frame, but rather show the raw number of false detections generated in approximately 54 minutes of data. Because each pixel of each frame has the potential to produce a false detection, in general, as frame rates are reduced, the number of possible false detections are reduced as a result. This property is reflected in the MWIR-1 band. The MWIR-2, not shown, exhibited a similar trend. For the LWIR, however, the opposite trend occurs and false detections increase as frame rates decreases. This phenomenon is likely due to the higher scene contrast of the LWIR sensor and the effect of scene motion. Scene motion at lower frame rates is more apparent from frame to frame. To test this theory, the LWIR and MWIR-1 data were re-analyzed using only files containing stationary scenes. For the stationary data set, the LWIR false detections decrease as the frame rate decreases, and is shown in Figure III.12. Though not shown in Figure III.12, false detections for dual LWIR/MWIR-1 were analyzed. The dual band false detections did not vary with frame rate and showed only minor random fluctuations.

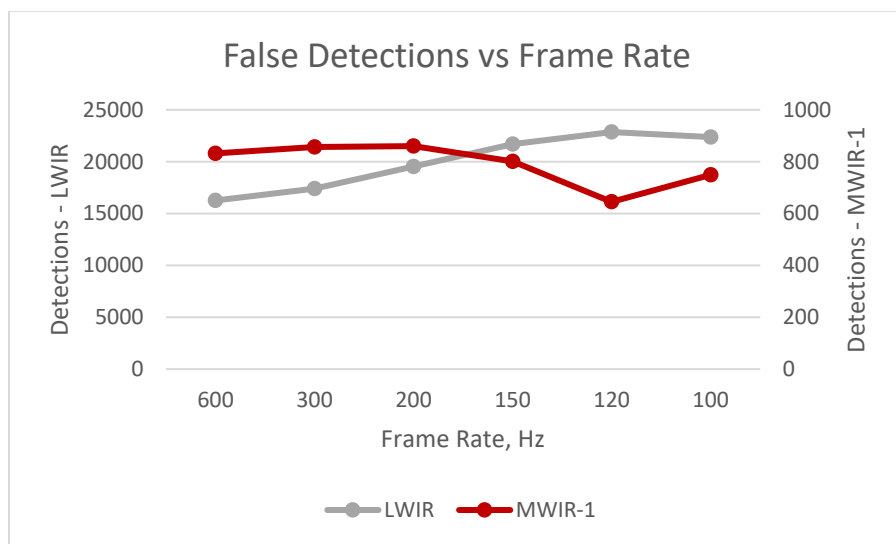


Figure III.11. False detections as a function of frame rate for LWIR and MWIR-1.

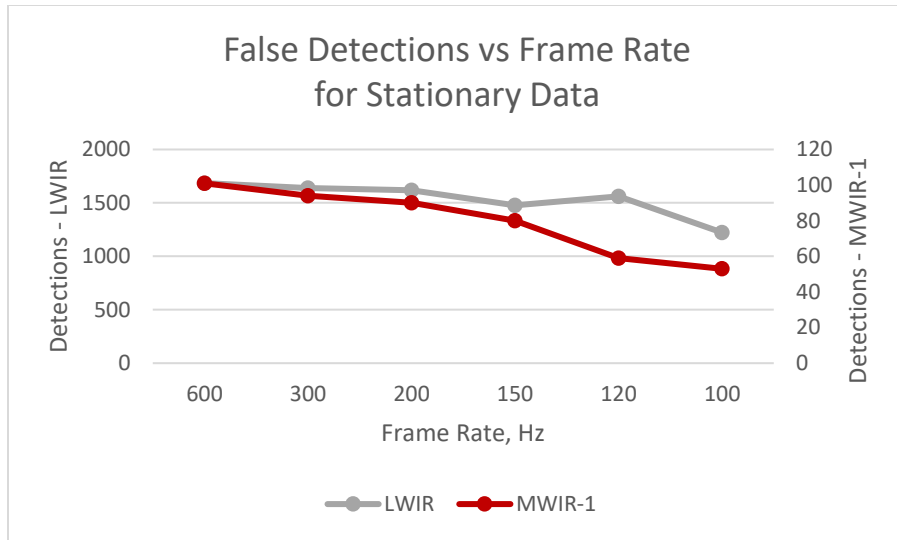


Figure III.12. False detections as a function of frame rate for LWIR and MWIR-1 for stationary data.

Classification

For classification, the results are presented as the output of the complete algorithm. Hostile fire events are first detected and then classified. In this way, the results of the classification scheme, dynamic distance, is limited by the detection rate and provides a representation of the system performance. The classifiers used in the analysis are linear discriminant analysis (LDA) and support vector machine (SVM). The classifiers were generated using temporal profiles at the ground truthed pixel locations of hostile fire events. The classifiers were evaluated using leave-one-out cross-validation. The classifiers were trained using as many as 13 features, denoted as the 10% width, the 50% width, fall time, and the 10 temporally/intensity normalized samples, to discriminate between CLASS 1, CLASS 2, and CLASS 3. For the case of dual band classification, up to 13 features were extracted from each band, providing for a total of 26 possible features. It is noted that these features are also used in the detection filtering process, and thus do not require additional processing to generate. The performance of the classifiers, with respect to differentiating hostile fire classes, using ground

truthed data, is shown in Table III.1. Here it can be observed that the hostile fire signatures are more distinguishable in the LWIR band, given the chosen features and the hostile fire data set. The CLASS 1 class is more easily classified, whereas the CLASS 2 and CLASS 3 class have more overlap in the LWIR band. Though reduced in performance, the MWIR follows a similar trend. The lowest performing band for classification, the MWIR-2, interestingly enough, has the alternate characteristic of differentiating the CLASS 2 and CLASS 3 classes well, with reduced performance between the CLASS 1 and CLASS 2 classes. The dual band LWIR/MWIR-1 classifier has increased performance over the individual bands. Unlike the dual band detection process, classification is not limited to the worst performing individual band. Table III.2 provides similar classifier performance results for 600 Hz data. The actual number of events of ground truth data for each band varied slightly, introducing some variation in the classification results. The trends in performance, however, are representative of the data set as a whole.

Table III.1. Classifier performance at 120 Hz per band, using groundtruth data.

<i>Classifier Performance, 120 Hz</i>				
<i>Threat / Band</i>	MWIR-2	MWIR-1	LWIR	Dual LWIR/MWIR-1
<i>CLASS 1/Other</i>	83.69%	95.29%	97.39%	98.84%
<i>CLASS 2/CLASS 3</i>	96.44%	87.64%	94.44%	94.91%

Table III.2. Classifier performance at 600 Hz, using groundtruth data.

<i>Classifier Performance, 600 Hz</i>				
<i>Threat / Band</i>	MWIR-2	MWIR-1	LWIR	Dual LWIR/MWIR-1
<i>CLASS 1/Other</i>	94.26%	97.69%	99.24%	99.19%
<i>CLASS 2/CLASS 3</i>	98.55%	93.15%	94.81%	98.55%

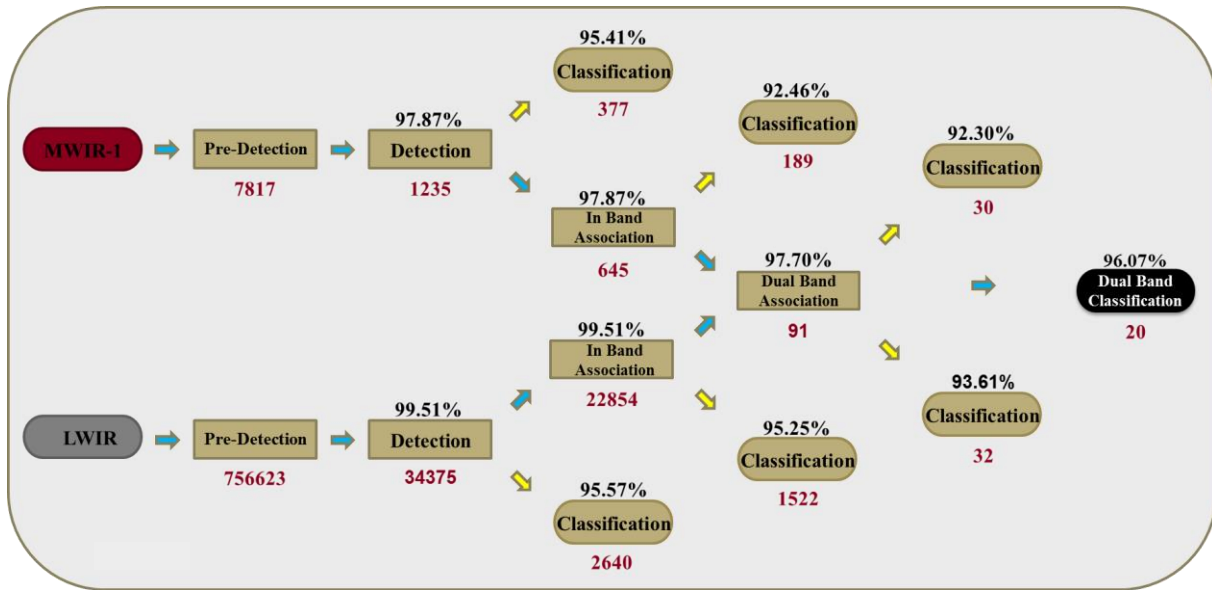


Figure III.13. 120 Hz data. Probability of hostile fire declaration (black percentage) and false alarms (red numbers) for various architecture configurations, including dual band detection and classification. Classification is performed using the dynamic distance classification scheme.

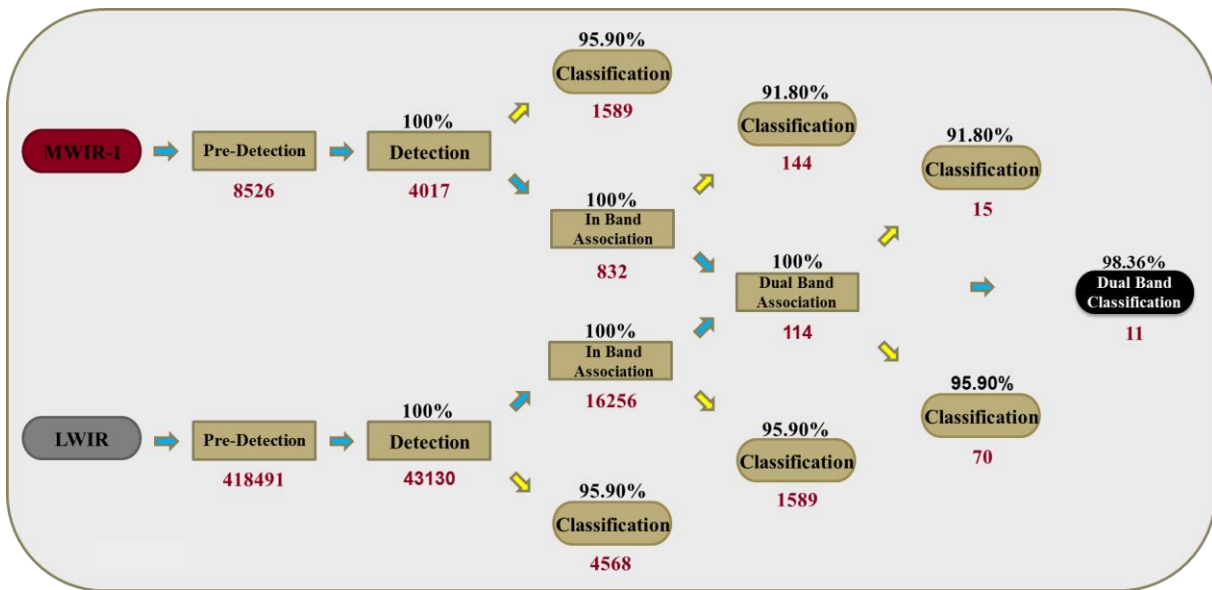


Figure III.14. 600 Hz data. Probability of hostile fire declaration (black percentage) and false alarms (red numbers) for various architecture configurations including dual band detection and classification. Classification is performed using the dynamic distance classification scheme.

System performance is shown in Figure III.13 for 120 Hz data. In this diagram each stage of the architecture displays the probability of hostile fire declaration and number false alarms generated at the stage. Here, hostile fire declaration is the result of correctly detecting a hostile

fire event and classifying the detection as any of the hostile fire classes. This is a fair representation of system performance given that threat class classification is a secondary objective to the threshold requirement of declaring hostile fire. False alarms are non-hostile fire detections, which have been classified as hostile fire. The diagram in Figure III.13 shows the results of various architecture configurations. For example, the dual band system requiring both LWIR and MWIR-1 detections, employing the dynamic distance classification scheme using features from both bands, has 96.07% probability of declaring hostile fire. Alternately, a dual band system which does not employ a classification scheme, has a higher probability of hostile fire declaration at 97.7%, but also has 91 false alarms. Single band systems for LWIR and MWIR-1 are also visualized. Figure III.14 provides an equivalent system performance analysis for 600 Hz operation using LWIR and MWIR-1. For comparison, single band MWIR-2 performance at 600 Hz and 120 Hz is presented in Figure III.15.

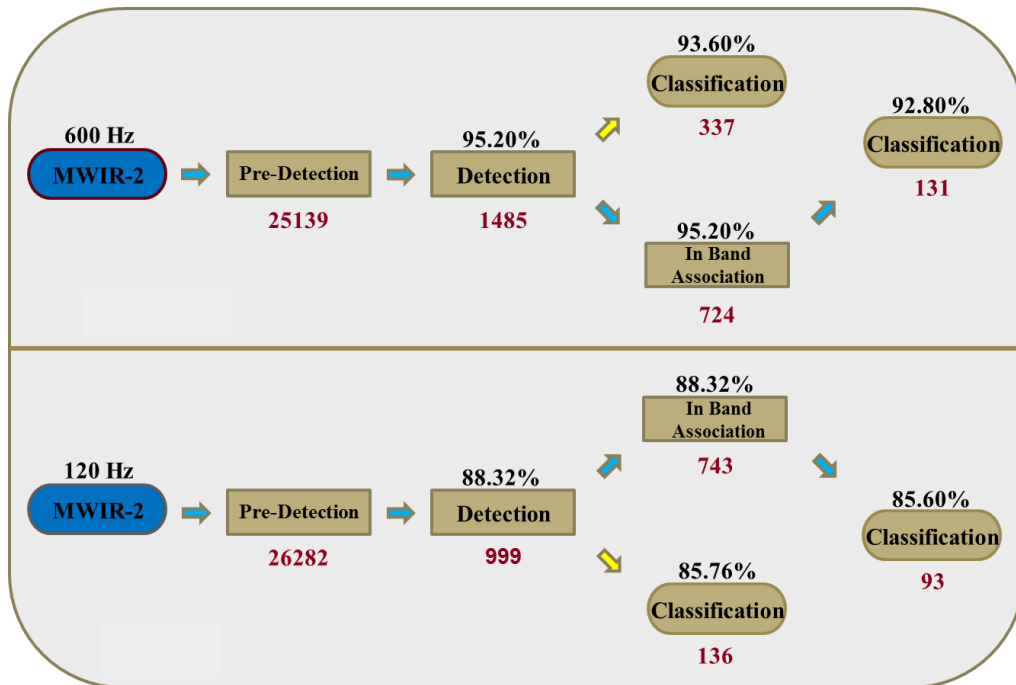


Figure III.15. 600 Hz and 120 Hz MWIR-2 data. Probability of hostile fire declaration and false alarms for various architecture configurations for single band detection and classification. Classification is performed using the dynamic distance classification scheme.

Figure III.13 and Figure III.14 provide valuable insight into system performance. First, the choice of applying an in-band association filter which selects the best detection candidate within a spatial-temporal volume, reduces potential false alarms without degradation in detection rate. Additionally, the false alarm rate is significantly reduced when the best candidate detections in the individual bands are associated across bands within a spatial-temporal volume. Thus, the exceptional number of false alarms generated in the LWIR due to scene motion are removed when compared against the relatively featureless imagery of the MWIR-1. The false alarms displayed in Figure III.13, Figure III.14, and Figure III.15 are from a subset of the false alarm data set which included data collected on a sunny day. It was found that this data was the most challenging for the system, and therefore, is the most relevant. The false alarm data is influenced by two variables, weather conditions and scene motion, which affect system performance. The false alarm results are shown as a function of these two variables in Table III.3 and Table III.4, where the values have been scaled to represent false alarms per hour. The sunny data more dramatically affects the MWIR bands. The LWIR band false alarms are greatly increased when motion occurs. The dual band system minimizes the effects of scene motion and weather condition, but more work is needed to reduce false alarms to a practical number.

Table III.3. False detections and false alarms per band as a function of weather condition for 120 Hz operation.

<i>False Alarms per Hour by Weather Condition</i>								
<i>Band</i>	MWIR-2		MWIR-1		LWIR		Dual LWIR/MWIR-1	
<i>Day</i>	Cloudy	Sunny	Cloudy	Sunny	Cloudy	Sunny	Cloudy	Sunny
<i>False Detection</i>	36.1	825.6	54.8	716.7	15076	25393	5.2	101.1
<i>False Alarm</i>	0.0	103.3	23.2	210.0	813.5	1691.1	1.3	22.2
Note: Day Cloudy 93 min, Day Sunny 54 min								

Table III.4. False detections and false alarms per band as a function of scene motion for 120 Hz operation.

<i>False Alarms per Hour by Scene Motion Type</i>								
<i>Band</i>	MWIR-2		MWIR-1		LWIR		Dual LWIR/MWIR-1	
<i>Scene Motion</i>	Static	Moving	Static	Moving	Static	Moving	Static	Moving
<i>False Detection</i>	185.0	706.2	32.8	1032.3	877.2	68681	2.2	146.2
<i>False Alarm</i>	14.4	103.1	14.4	306.2	50.6	4141.5	0.0	33.8
Note: Static Scene 108 min, Moving Scene 39 min								

The resulting hostile fire declaration in Figure III.13 can be further decomposed to provide a more complete view of system performance in Table III.5. In Figure III.13, the dual band detection shows 2.3% of hostile fire events were not detected. Table III.5 indicates that a greater percentage of CLASS 3 targets were undetected, when compared to the other classes. Similarly, the hostile fire events mistakenly classified as false alarms predominantly affected the CLASS 3 class. Here, it is observed that the majority of hostile fire events are detected and declared as hostile fire, yet there is strong overlap with the CLASS 2 and CLASS 3 classes. This result is expected, given the classifier results using ground truthed data in Table III.1. The reduced performance for CLASS 3s can be attributed to the increased variability in CLASS 3 signatures, and factors relating to increased range. Increased ranges can make CLASS 3 signatures more susceptible to atmospheric degradation and more vulnerable to the effects of poor focus. As signals are increasingly attenuated, the signatures gradually lose their characteristic shape. The effects of attenuation and blurring require further evaluation to determine the limits of reliable classification and detection.

Table III.5. Confusion matrix for dual band LWIR/MWIR-1 at 120 Hz. Undetected hostile fire events are also included.

Confusion Matrix					
Dual Band Classifier 120 Hz					
<i>Stim</i> \ <i>Resp</i>	<i>Class 1</i>	<i>Class 2</i>	<i>Class 3</i>	<i>FA</i>	<i>Not Detected</i>
<i>Class 1</i>	95%	2%	0%	0%	2%
<i>Class 2</i>	3%	96%	0%	1%	0%
<i>Class 3</i>	0%	15%	74%	5%	5%

Real-Time Operation

Processing time, or real-time performance, is an additional metric which should be considered when evaluating algorithm performance. An initial analyses showed that the number of pre-detections which are recorded when processing on-the-move data can become exceptionally large. It was found that the temporal frequency filtering n –window system for storing pre-detections, along with the local spatial max pre-track selection and elimination process, tremendously reduce the number of pre-detections which are actually analyzed. There is a reduction between 92% and 97% in pre-tracks initialized versus pre-tracks analyzed, depending on band. These reductions still allow a large number of pre-detections to be analyzed, as shown in Figure III.13. Another potential mechanism for reducing the number of analyzed pre-detections is to increase the SNR_B threshold scaling value. An increase in SNR_B from 4 to 8 was shown to provide a 55% to 63% reduction in analyzed pre-detections. Though actual hostile fire events collected have SNR_B values above 8, it should be noted that these values are generated from stationary data. For on-the-move hostile fire data, it is expected that background clutter will increase, and reduce effective SNR_B for hostile fire. From Figure III.13, it can also be observed that the MWIR-1 pre-detections are significantly fewer than the LWIR pre-detections. Given a strategy which requires detection in both LWIR and MWIR-1, one potential option is to use the

MWIR-1 as a guard band. Therefore, the LWIR band is only evaluated after the MWIR-1 band reports a detection. In a worst-case analysis, where each detection in the MWIR-1 band, yielded a respective detection in the LWIR band, there was a 93% reduction in processed detections. The algorithm, developed using MATLAB, is currently being ported to C++ and OpenCV. The data was processed using a Xeon E5-2620 v3 12 core processor. The video data was 120 Hz with a frame size of 640x240. Real data was replicated to provide the extended rows. The average time to process a frame containing detections was found to be 5.34 ms. For 120 Hz operation, there are approximately 8.33 ms between frames. Thus, the algorithm presented in this paper can be operated using dual band LWIR/MWIR-1 at 120 Hz in real time. Using the MWIR-1 band as guard band could increase processing efficiency, progressing towards real time operation for large format camera frame sizes.

Discussion

Hostile fire detection and location for ground vehicles requires a solution which is practically implementable given vehicle motion. For on-the-move operation using ground vehicles, however, an increased number of false alarms are generated due to scene motion. This source of false alarms affects all bands, but more deeply impacts the LWIR band. Sun glints are also a source of false alarms, affecting MWIR bands more severely, but also affecting the LWIR band. A practical hostile fire system for ground vehicles will need to detect hostile fire, address false alarms due to scene motion and solar glints, and process this data in real time to be applicable.

The algorithm presented in this paper provides an architecture that can be used for hostile fire detection systems. The algorithm provides false alarm mitigation strategies when using single band systems. A significant improvement is demonstrated when these strategies are

extended to a multi-band system. Though the multi-band system can consist of any number of guard band sensors and verification sensors, results of a dual band LWIR/MWIR-1 system are presented. These results show acceptable hostile fire declaration rates, where declaration is classification as a hostile fire event. The false alarms, reduced to a manageable rate, leave room for further improvement. A possible avenue for improvement is to extend the number of features used in classification. Additional features could also be used to improve hostile fire classification rates, especially with respect to the CLASS 3 class.

The algorithm uses multiple methods to improve efficiency for real time operation. The features used for classification are reused from the detection process. Additional classification features must be balanced against performance improvement and processing degradation. Other improvements to processing time can be made from using the n -window system, increasing the SNR_B threshold, and using the MWIR-1 as a guard band. The algorithm was implemented, without the use of a guard band, as a dual band LWIR/MWIR-1 system for 640x240 sensors at 120 Hz in real time.

The algorithm is shown to be a potentially viable solution to on-the-move hostile fire detection. Improvements can still be made, especially for classification and false alarm mitigation. Though the detection and classification processes are separable, improvements to the detection process, such as additional detection filters, can potentially improve results as well. Future work will investigate additional detection filters and classification features. Additionally, an investigation into saturated signature profiles will be made. Though the algorithm can handle partial saturation of signatures, it is not intended to detect and classify heavily saturated signals. Moreover, techniques will be investigated to yield real time processing for large format sensors operating at 120 Hz.

IV. DETECTION, CLASSIFICATION, AND LOCALIZATION SYSTEM AND METHODOLOGY

Introduction

This chapter provides a generalized description of the algorithm developed in Chapters 2 and 3, and also provides multiple system designs with accompanying algorithmic approaches to overcome limitations of the impulse or weapon fire detection algorithm and emphasize its strengths. System designs include methods for minimization of saturation clutter and minimization of clutter using additional acoustic sensors. In one design, the methodology is layered to account for various levels of hardware availability.

The generalized method presented here is referenced in subsequent system designs and methodologies, when referring to the weapon fire algorithm, and in particular, when referring to weapon fire detection, and also weapon fire classification. The generalized method described in this chapter has been published by the United States Patent and Trademark Office under patent number US 10,389,928 B2 on August 20, 2019.

The chapter also provides a generalized EO/IR and acoustic weapon fire detection system design. The design does not emphasize saturation mitigation, but emphasizes clutter mitigation by using the acoustic sensor as a confirmation sensor to the EO/IR detection sensor. This system design provides for increased performance with minimal additional hardware and associated processing. The acoustic clutter mitigation system and method described in this chapter has been filed for application at the United States Patent and Trademark Office on August 9, 2018.

In addition, this chapter provides a description of multiple system designs which offer approaches to reduce the effects of saturation, and also further minimize clutter. Though the system designs address similar problems, differences among designs lay in required hardware, affecting size, weight, power, and cost. Additionally, there are performance differences. For

example, a system using an acoustic detection sensor and integrated on a ground vehicle, can exhibit lower range performance when compared to optical detection sensors. The acoustic sensor also has a latency associated with the rate of travel for sound waves. A system design using GPS assumes that GPS information is available. The saturation mitigation systems described in this chapter have been published by the United States Patent and Trademark Office under patent number US 10,209,343 B1 on February 19, 2019.

And finally, this chapter provides an in-depth procedure for utilizing EO/IR imager, real time clock, compass, gyroscope (gyro), and Global Positioning System (GPS) data to detect the sun. Sun detection is used as means of minimizing saturation clutter. Methods for detecting the sun are provided for various levels of hardware availability. For example, in some environments GPS data may not be available, or some systems may not be equipped with a gyro. In these cases, information can still be extracted to allow for sun detection and saturation mitigation.

The position of the sun can be referenced as a pixel location, or an azimuth and elevation, depending on the availability of input information. Beyond determining the position of the sun, the orientation heading can be updated to provide more precise measurement. Additionally, the sun can be tracked when it is no longer detected in the current frame, such as due to obstruction. The sun detection method presented here is intended to be used on a moving platform or vehicle. The same method can be employed at night, during which the moon is detected, rather than the sun

Background

Weapon fire detection and localization can be accomplished using a variety of sensing modalities. Typical modalities include radar, acoustic, electro-optic, or some combination of

these. Electro-optical solutions typically exploit projectile launch blast, thermal radiation of in-flight round, and the thermal radiation of rocket motors of missiles and rockets. Electro-optical solutions are often employed for close range small arms and unguided projectile detection. Additionally, air-systems have employed electro-optical sensors for long range weapon fire detection systems such as missile warning systems. For systems detecting weapon fired air-to-air or air-to-ground, atmospheric signal degradation is less impactful than for systems which detect ground-to-ground weapon fire. Long range electro-optical weapon fire detection systems for ground application must be able to handle complications due to atmospheric signal degradation and near field clutter sources.

Electro-optical weapon fire detection systems for ground application, capable of detecting short and long range threats, require complex detection methodologies to accurately detect weapon fire over a broad dynamic range of signal intensity. When a high intensity weapon signature is available, as when the launch is near the sensor, simple thresholds can be used to detect weapon fire while minimizing clutter sources. Long range electro-optical weapon fire detections often involve severely reduced signature intensities due to atmospheric degradation. Additionally, near field motion can generate clutter sources of similar intensities to far field weapon fire. A complex detection algorithm is required to differentiate actual weapon fire signatures of various intensities from cluttered backgrounds while providing acceptable detection rates.

Acoustic weapon fire detection systems are typically used for small arms detection. Acoustic systems rely on the supersonic speed of the passing weapon fire projectile to produce a sonic boom (or “crack”). Additionally the acoustic sensor detects the sound wave (“bang”) generated by the weapon fire muzzle blast. The delay between the “crack and bang” is used to

estimate the range to the weapon fire. Multiple acoustic sensors are used to determine the bearing of the weapon fire projectile.

Sun positioning algorithms and sun detection and tracking systems are commonly used in not only space craft applications for navigation, but also for solar power systems and unmanned vehicle systems. As solar power systems have progressed to be a feasible solution for renewable energy, the need to track the trajectory of the sun accurately has amplified in order to provide increased yield. The sun tracking systems are typically considered as either closed-loop or open loop. For closed loop systems, sensors provide feedback to a controller, which adjusts an output such as the pointing angle of collection sensor. An open-loop system does not use feedback, and its output is determined by an algorithm and the current state of the system. An example of an open-loop system uses date, time, longitude, latitude, and solar positioning algorithm to compute the direction of the solar vector. Open-loop systems are often simpler and less expensive than closed-loop systems, but cannot correct for errors or anomalies. Closed-loop sun detection and tracking systems use dedicated sun detection sensors to determine the location of the sun. In this chapter, sun detection and tracking is performed using multipurpose focal plane arrays in combination with detection algorithms. The detections can be confirmed and credited to a solar direction vector using solar positioning algorithms and current date and time, reducing the occurrence of false sun detections. The work presented here is applicable to mobile platforms as well as stationary platforms. EO/IR imagers are currently used in many applications for surveillance and navigation. This concept provides increased capability for those application without the need for increased, dedicated hardware. For these applications, the sun can be determined within the field of view of the EO/IR imager.

Problem Description

A methodology for practical implementation of weapon fire detection. This method provides a framework for designing a hostile fire detection system. The method or algorithm is independent of sensor waveband and is also independent of the number of sensors. The algorithm emphasizes techniques for not only detection of weapon fire events, but also techniques for rejection of clutter as well. For on-the-move applications such as for a vehicle, clutter sources may not be consistent or predictable, and therefore an adaptive algorithm is provided to adjust to a changing scene environment.

System/algorithm modifications incorporating acoustic sensors to improve capability and increase feasibility are presented. Current EO/IR algorithms provide high detection capability performance but require some improvement to clutter rejection. This system provides a method to increase clutter rejection, as well as providing the additional functionality of range estimation. Other systems have used an Acoustic - EO/IR method. Here we argue that the particular EO/IR method benefits from the addition of limited, minimal acoustic hardware and processing.

System/algorithm designs to overcome limitations of the impulse or weapon fire detection algorithm and emphasize its strengths are presented here. The additional system components and algorithms address increased clutter rejection and saturation. Excessive false alarms from clutter can make the system unusable for practical implementation. The effects of saturation can limit detection capability and reduce the ability to distinguish between real saturated targets and saturated clutter targets.

For EO/IR sensors, the sun can often cause saturation, removing distinguishing scene information used in image processing. By being able to detect the sun, the resulting saturated pixels, as well as the source of saturation, are determined. These pixels, being known to have

been saturated by the sun, can be masked, ignored, or potentially be adaptively processed, thereby limiting confusion between other sources of saturation.

Sun detection systems often require dedicated sensor hardware to perform the detection and tracking of the sun. Sun positioning algorithms can also be employed which determine the direction of the sun based on inputs, such as current date, time, and location. The orientation of the system platform is still required to determine the relationship of the sun with respect to the sensing system platform. However, for applications which attempt to reduce hardware and minimize size, weight, power, and cost, adding additional sensors may not be a preferred solution. EO/IR sensors are currently used in many applications for surveillance and navigation. The in-depth methodology provides sun detection and tracking capability, using EO/IR sensors only. The detected sun is used to identify sources of saturation within the sensor's field of view. In some scenarios, GPS may not be accessible and a compass may not be available. Additional information, such as orientation heading can be provided by incorporating sun positioning algorithms, current date and time. With the addition of a gyro, the sun location can be estimated or tracked from frame to frame. This step can also be performed using image processing of scene motion. Sun tracking is critical, as the vehicle motion can often cause the sun to be fully or partially obscured by near range objects, resulting in inconsistent detections. By enabling sun tracking, predetermined detections can be maintained even during instances of obstruction. Thus "flickers" of sun, partially obscured, can still be tracked and identified. Additionally, the inclusion of orientation data from a compass, can provide an estimate of the current time.

Methods

A generalized weapon fire detection method is presented here. The method generalizes the techniques described in chapters 2 and 3. The method described in Figure IV.1 is shown for a

single sensor, but can be extended to multiple sensor. Figure IV.2 and Figure IV.3 provide a more detailed description of the detection and classification process, respectively. The generalized detection method will be used in subsequent system designs presented in this chapter.

As an example of a multiple sensor system, a detection system utilizing three imaging sensors is proposed. The three imaging detection sensors, one imaging in the long-wavelength infrared (LWIR) band, and the other two, imaging in the mid-wavelength infrared (MWIR) band, operate at a frame rate of 120 Hz and have a focal plane array (FPA) size of 640 by 480. Here, the two MWIR bands are spectrally non-overlapping. The three detection sensors, contained in the weapon fire detection imaging sensor system, provide spectrally distinct signatures of weapon fire events. The extracted weapon fire signatures provide intensity, duration, and shape information used for determining detections. Similar to the two-band system, detections from each band are compared in both space and time to determine if they are a result of the same event. Additionally, by analyzing the spectral signature in three bands from temporally and spatially synchronized weapon fire event signatures, a significant amount of clutter can be rejected. This clutter rejection is critical for on-the-move ground vehicle application against ground weapon fire targets that can occur at both near and far tactical ranges. The use of intensity, shape, and duration measurements from multiple temporally and spatially aligned spectral signatures improves upon prior art, and allows for on-the-move application of weapon fire detection at relevant tactical ranges.

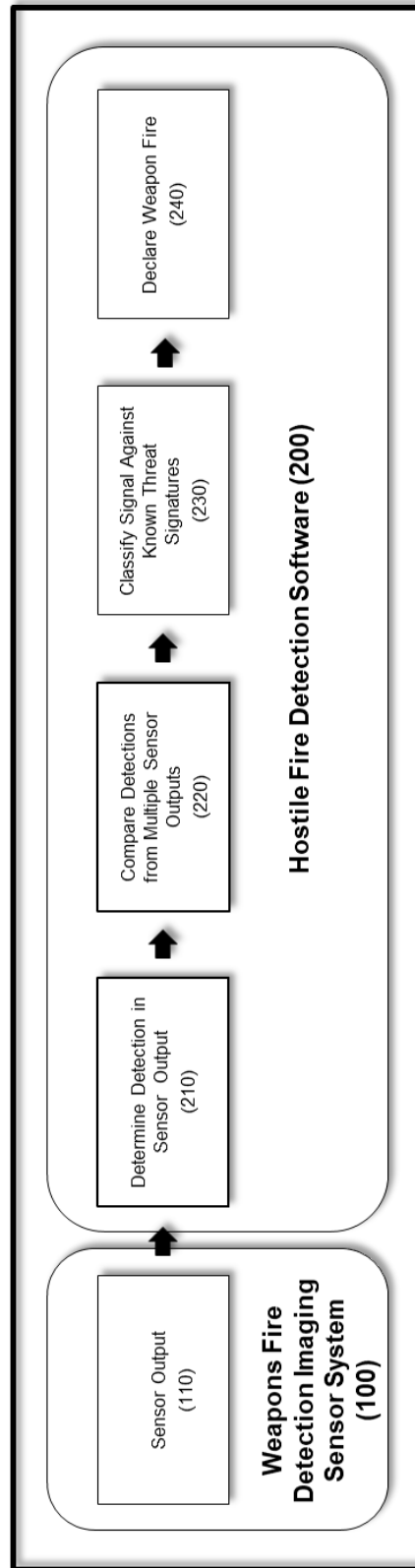


Figure IV.1. Detection Software Diagram

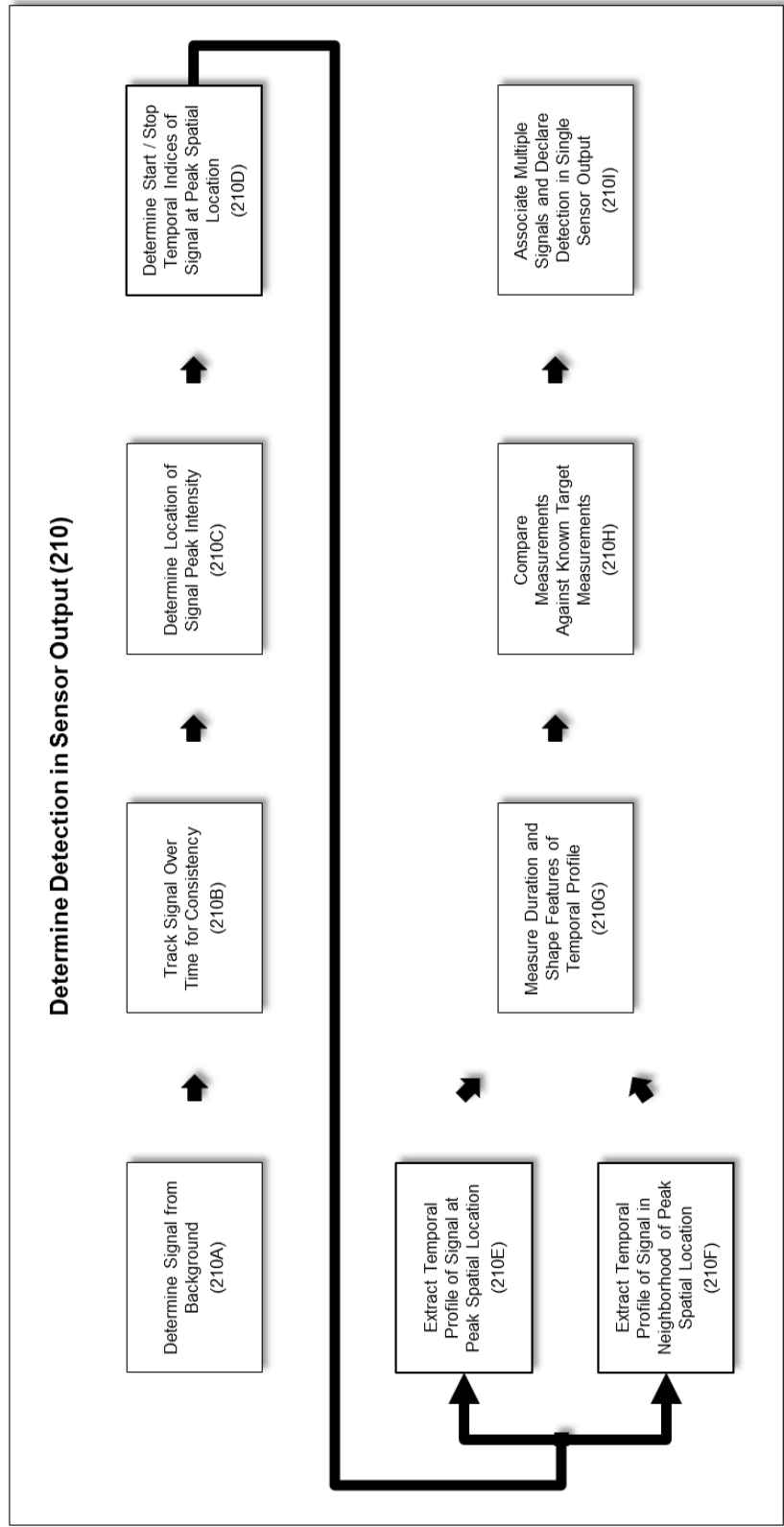


Figure IV.2 Detection decision diagram

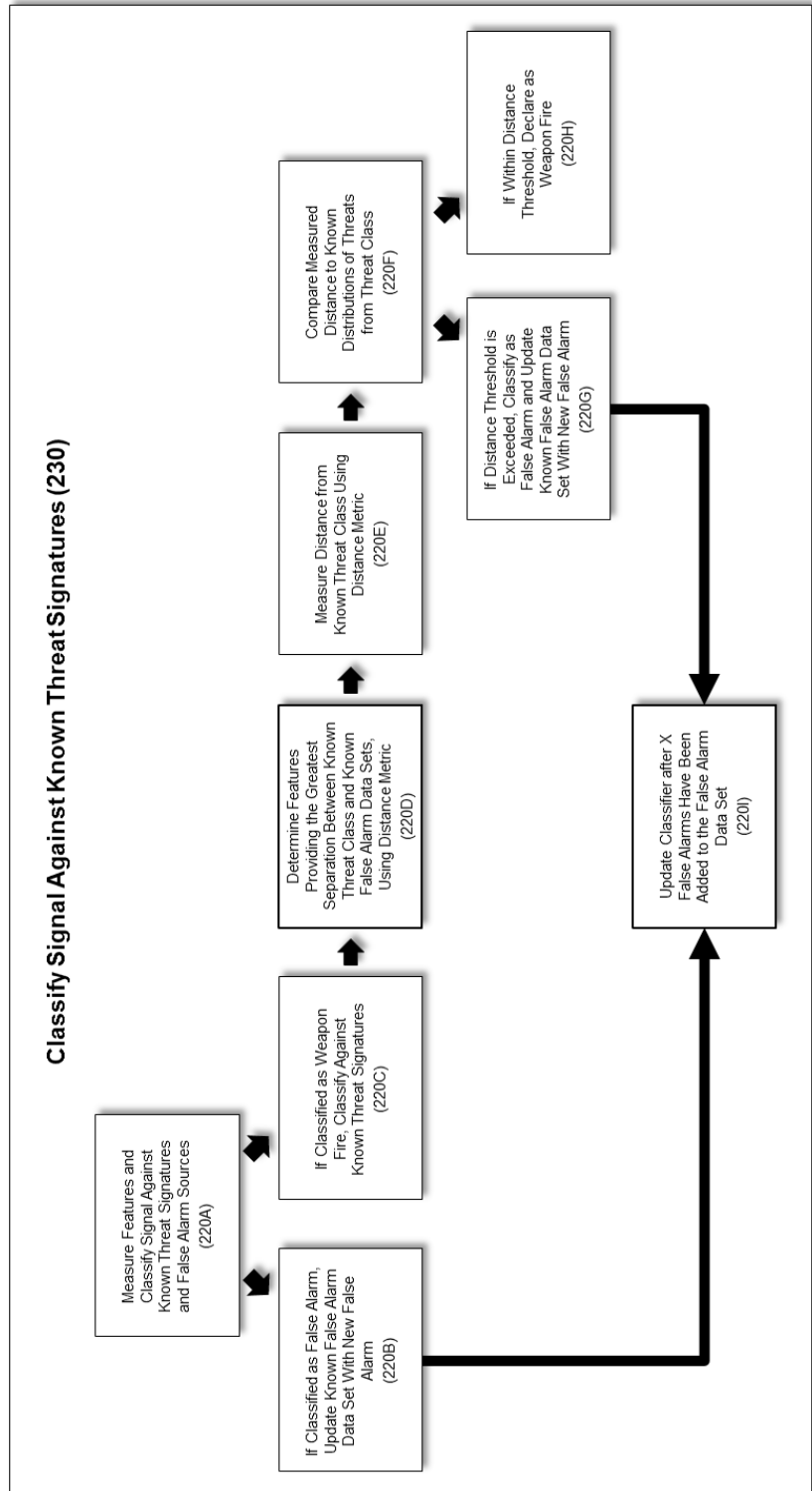


Figure IV.3. Classification decision diagram

The three-band weapon fire detection system improves upon the weapon fire detection system presented in the Chapters 2 and 3. The additional detection band increases the information available to be used for clutter discrimination. The reduced number of false alarms allows for the weapon fire detection systems to be used in challenging scenarios, such as strong sun and on-the-move operation (see Table III.3 and Table III.4). The increased number of spectral bands also reduces the opportunity for saturation clutter to occur across all three spectral bands. Saturation clutter occurs when a pixel location in each band, all being mapped to the same location in object space, is saturated for a given duration. Thus, the observation of an event generated a saturated signal across all spectral channels. A saturated signal, having an appropriately short duration, can be indistinguishable from the saturated signal observed from a weapon fire event, due to loss of information. The sun, and reflections thereof, can generate significant saturation clutter. This phenomenon often occurs during apparent scene motion, as the sun “flickers” behind partial obscurations. Initial empirical observations support the widening of the detection spectral bandwidth by adding independent detector spectral bands to reduce the chance of all bands being saturated during a given window of time, at a specified location. Though saturation clutter is not completely eliminated, the idea of adding additional sensors for the specific case of reducing saturation clutter is explored in subsequent designs.

There are disadvantages to adding additional imaging sensors, and in particular, relate to size, weight, power, and cost (SWAP-C). The addition of an imaging sensor increases system SWAP-C, due to the addition of the EO/IR detector and the accompanying optical system. For thermal imagers, the detector cost will increase, when compared to a visible camera. Even when low-cost thermal detectors are used, the required optical system can often become a cost driver.

Optical apertures for longer wavelengths (thermal) have increased diameter, when compared to shorter wavelengths (visible), to improve resolution. And thus, each independent aperture and optical system for each added spectral band can substantially increase system size. Though space can be saved using a single aperture and spectral beam splitters, cost is considerably increased due to the more complex optical system.

Similarly, to providing an additional optical sensor to improve clutter mitigation, an acoustic sensor can be added to the system design for clutter mitigation. Here the acoustic sensor is used as a weapon fire detection confirmation sensor to the imaging detection sensors. The acoustic sensor's region of interest must encompass the detection sensor system's region of interest. Given a detection within the imaging bands, the acoustic signal is scanned for an accompanying signature in the acoustic channel. Accurate time stamping of detection sensor system output and acoustic saturation sensor output will allow for detected signals from the same event to be correlated. Acoustic signals, travelling at the speed of sound, will arrive at the detector later than optical signals, traveling at the speed of light, such that the arrival time of the acoustic signal is impacted due to target range and atmospheric conditions. Ambient temperature data can be used to provide a better estimate the speed of sound. Multipath delayed detections must also be taken into consideration for accurate processing. In practice, close range threats need to be detected as soon as possible, for potential action to be taken, and therefore the detection delay attributed to an acoustic sensor may not be acceptable. The acoustic sensor, in combination with the optical sensor, can be used to reduce clutter, estimate an approximate range to target, and can potentially offer lower SWAP-C, lower bandwidth, and reduced processing requirements.

Similar to traditional clutter mitigation techniques, saturation mitigation solutions can be achieved through multiple approaches. One potential hardware mitigating solution to the saturation problem is to increase the dynamic range of the thermal sensor, which effectively reduces the maximum range of sensor to target which yields a saturated signal. For many applications, the constraints of size, weight, power, and cost (SWaP-C) restrict the availability of sensors with higher dynamic range. Saturation mitigation can also be achieved through system design, which incorporates hardware and software approaches.

For many applications, high SWaP-C sensors may not be feasible, however, additional low SWaP-C sensors may be an approach for consideration. For instance, a saturation sensor can be added to a detection system for the verification of weapon fire events which saturate the detection system. The saturation sensor can be an optical sensor with a limited effective aperture, to reduce the amount energy which reaches the FPA during saturation events. Alternatively, a saturation sensor with reduced sensitivity, can increase the fluence at the aperture required to saturate. For detection of distant events, the detection system requires high sensitivity, and narrow instantaneous field of view (iFOV), to increase signal to noise for events which may subtend a small angle. For persistent observation of a large field of view, this is typically achieved with larger format sensors, thus increasing the size of the FPA, associated optics, and SWaP-C. The optical saturation sensor, however, is specifically addressing close range weapon fire which exhibits signals with high intensity and a larger spatial extent. The saturation sensor is also functioning as a verification sensor with event detection occurring from the detection sensor. For these reasons, a saturation sensor could be configured using a smaller FPA format and, smaller optics. The optical saturation sensor and detection system should be aligned and registered, such that the saturation sensor's region of interest encompasses the detection sensor's

region of interest. Additionally, signals generated from the saturation sensor must be temporally registered to signals generated from the detection sensor. There are many viable options within the trade space of optical and temporal alignment and resolution tolerances. For example, a high fidelity temporal alignment or frame time stamping may allow for lower tolerances in optical alignment. Thus, a saturation sensor with a low FPA size can be used such that each pixel in the saturation sensor, is registered to a region within the detection system sensor(s). The saturation sensor can be extended to a system of saturation sensors.

The use of imaging sensors for weapon fire detection have some advantages in comparison to traditional acoustic weapon fire systems. The signal travels at the speed of light, and thus the time of launch can be determined, and the detection can be made with minimal latency. Additionally, the spatial location, or azimuth and elevation, can be determined from the detection pixel location. The subsequent system design incorporates an acoustic sensor for the special case of saturation clutter.

The acoustic saturation sensor, similar to the optical saturation sensor, allows for signal extraction when the imaging detection sensors are saturated. This signal extraction provides for improved detection of close-range threats that would otherwise saturate the detection sensors, and also allows for removal of saturation clutter.

The acoustic saturation sensor and weapon fire detection sensor operate on the assumption that weapon fire within close enough proximity, and having enough energy, to saturate the imaging detection sensors, will generate an audible and detectable sound. In contrast, solar clutter, capable of saturating the imaging detection sensors, will not generate an audible sound. In this case, the imaged signature is saturated and not recovered. The saturation is a cue, however, to the potential occurrence of an event. The event is confirmed as a weapon fire event,

when the saturated signal is followed by an audible signature having an intensity, duration and shape profile matching that of known weapon fire events. Here, acoustic signatures are used as a confirmation, rather than being a means of informing a detection. With knowledge of tactical weapon ranges, ranges at which these weapons saturate the imaging detection sensors, and an approximation of the speed of sound, a narrow search window can be used, reducing potential acoustic clutter.

Traditional acoustic weapon fire detection systems perform detection and localization (azimuth, elevation, and range) via acoustic sensors only. This requires multiple microphones needed to estimate direction of travel, weapon type, and velocity estimation using signals generated from both the shock wave front and the sound wave front, also known as the “crack and bang.” The crack and bang detection method is used for supersonic threats, which generate a shock wave that is detectable when the projectile passes within range of the acoustic detection sensor. The design concept presented here is capable of detecting both subsonic and supersonic threats, and also does not require the projectile to pass within range of the detection sensor. Because the imaging detection sensor systems is able to perform initial detection, azimuth localization, elevation localization, and time of launch, a reduction in complexity for the acoustic system is able to be achieved. Range estimation to weapon fire launch is also readily achieved using a combination of imaging and acoustic sensors. An acoustic saturation sensor has reduced SWAP-C when compared to an imaging saturation sensor. A disadvantage of acoustic systems is the inherent latency caused by the delayed arrival of sound waves, when compared to optical systems.

An additional weapon fire system design is presented in Figure IV.4. The design includes a detection system which incorporates position, orientation, and timing data. Here, the depicted

detection system is a three-band detection system, but could utilize any number of detection sensors. Position, orientation, and timing data can be provided from an inertial measurement unit and GPS receiver. A local clock and additional sensors could potentially provide position, orientation, and timing data if GPS data were not available. Given position, orientation, date, and time, the position of the sun, relative to the detection system, can be calculated using ephemeris data.

The premise for the weapon fire system design using position, orientation, and timing data, is that the additional information allows for increased saturation clutter rejection due to the known position of the sun, and provides for improved detection of close range weapon fire events. The sun, being a strong source of saturation clutter, is able to be masked, if its position within the sensors field of view is known. The assumption here is that other sources of saturation will be from actual weapon fire events. An advantage of sun detection, is that it is able to remove a large clutter source without the inherent latency, as seen by an acoustic system. This method attempts to increase performance by increasing detection of true negatives, and thus reducing type II error (false alarm rate), rather than increasing detection of true positives, which would reduce type I error (miss rate). A disadvantage is that this method assumes that position, orientation, and time data are available, or can be approximated, which may not be the case in all

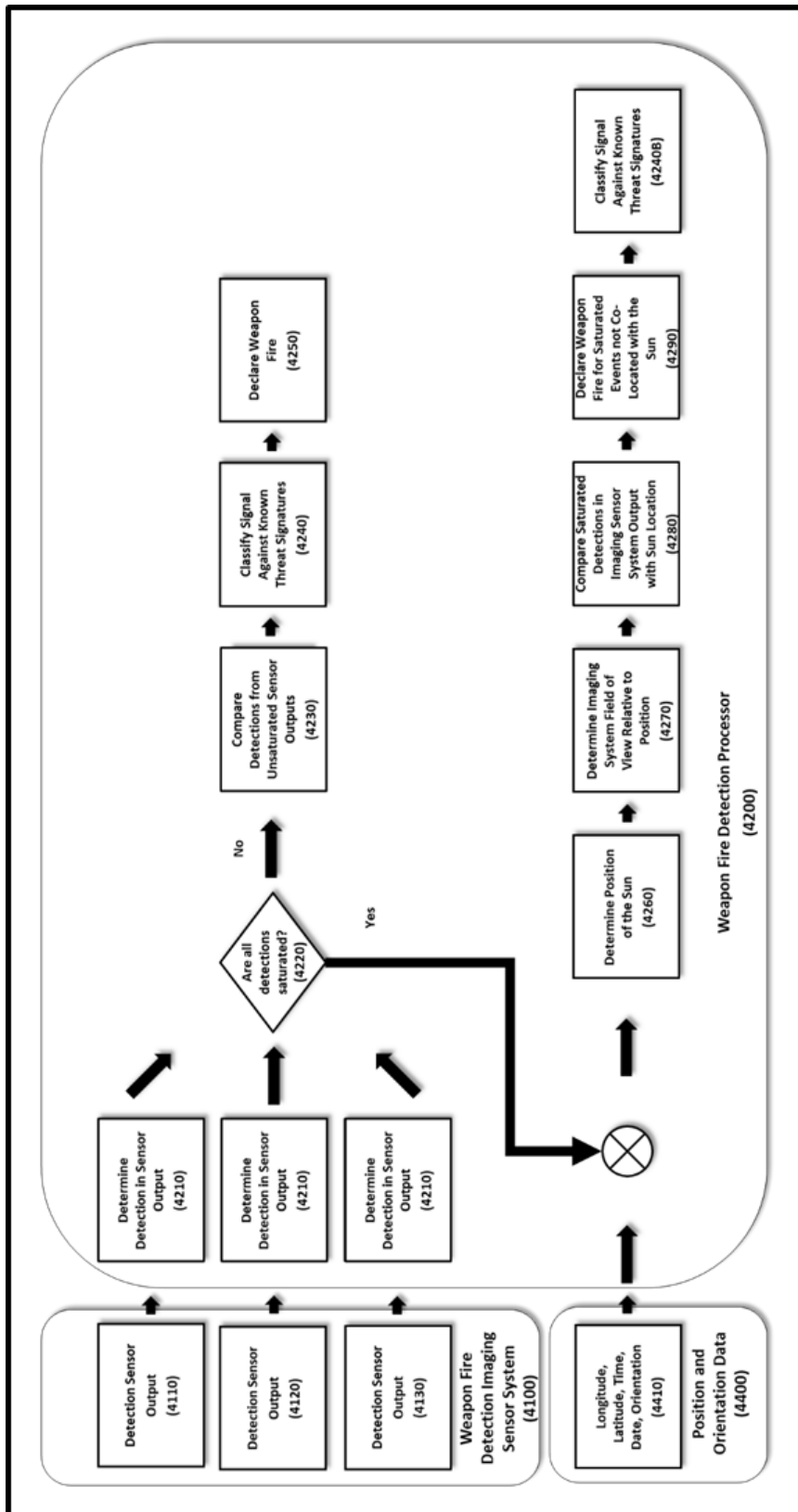


Figure IV.4. Weapon fire detection system processing using GPS for sun detection

environments. However, if the position, orientation, and time data are available, significant additional hardware is not needed, but only the ability to process the additional information, allowing for improved performance with only minimal increase in SWAP-C, without sacrificing latency.

The method of sun detection can be further improved using additional algorithmic approaches. The sun detection method presented is a deterministic approach based on provided inputs of position, orientation, date, time, and ephemeris data. These inputs may include errors due to the sensors providing data to the system. Position estimation error may not significantly contribute to errors in the sun's relative position, but orientation error can potentially provide noticeable error in the sun's relative position. Date and time accuracy to within minutes will not generate noticeable error for this application. Sun position estimates can be improved, however, through image processing techniques. Given an estimate of the sun's position, a search window can be applied within the image frame, encompassing the position estimate. The actual position of the sun can be determined by searching within the window. One detection method that can be applied is the detection of persistent saturation. Once the position is determined in the current frame, the position can be estimated in the next frame based on inertial measurement data or global scene motion analysis. Estimation of the sun's position in future frames can be employed during scenarios when the sun becomes obscured. The incorporation of image processing techniques adds additional processing but may allow robustness for various levels of data availability and accuracy.

The system depicted in Figure IV.5 allows for sun detection using various levels of hardware availability and image processing techniques. The sun detection process is described in Figure IV.6 and provides methods for detecting, locating, and tracking the sun. The method

allows for determination of the sun’s location using image processing techniques and further confirmation of the sun’s location if real time clock and orientation data are available. In Figure IV.7, an image processing based method to determine the sun’s location within the image is presented. The sun can then be tracked using either image processing techniques or angular velocity data from a gyro, as shown in Figure IV.8. The sun orientation can be determined using various levels of hardware availability. In Figure IV.9, the sun’s orientation can be provided, given access to a real time clock or a compass. Accuracy can further be improved with access to position data, such as from a GPS receiver. Given the detected sun within the image and predicted sun location using ephemeris data, the orientation of the imaging system, and subsequently, the platform which it is mounted on, can be determined or improved upon. Additionally, persistent tracking of the sun can provide an estimate of time, given a known orbit determined by ephemeris data.

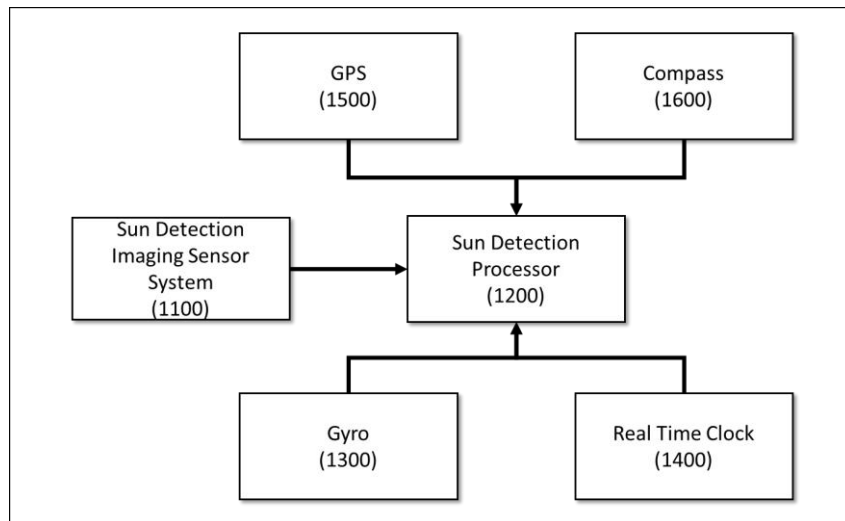


Figure IV.5. Sun Detection System

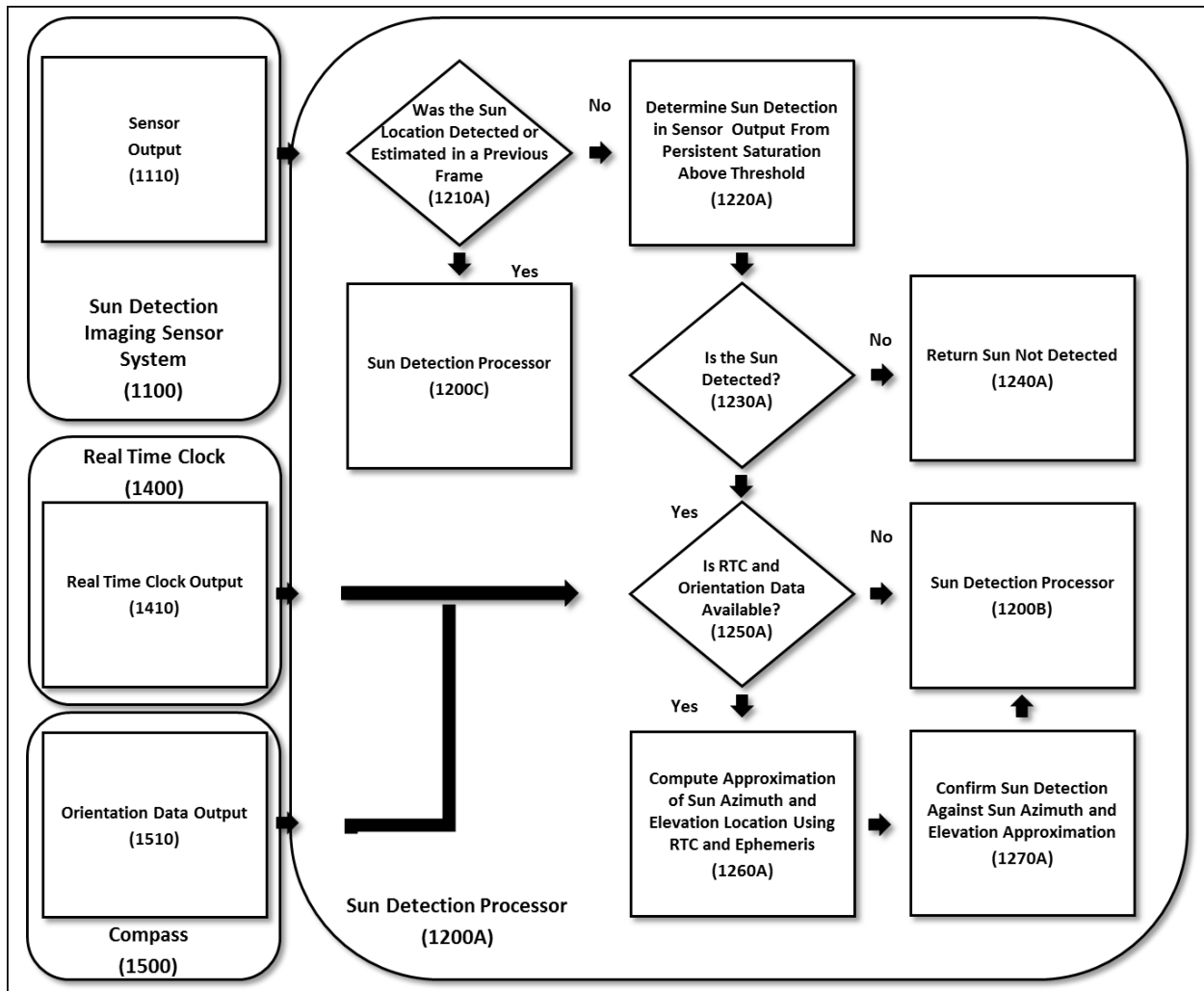


Figure IV.6. Sun Detection

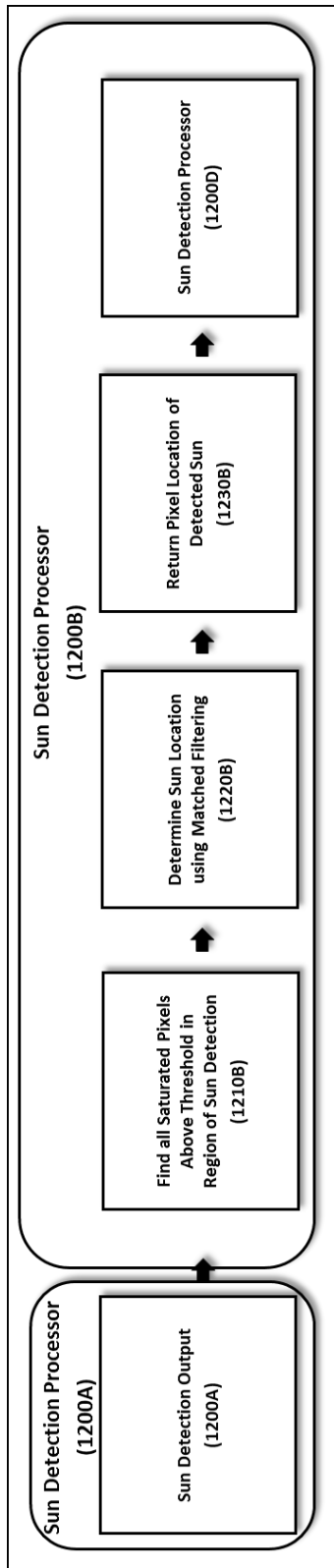


Figure IV.7. Sun Locating

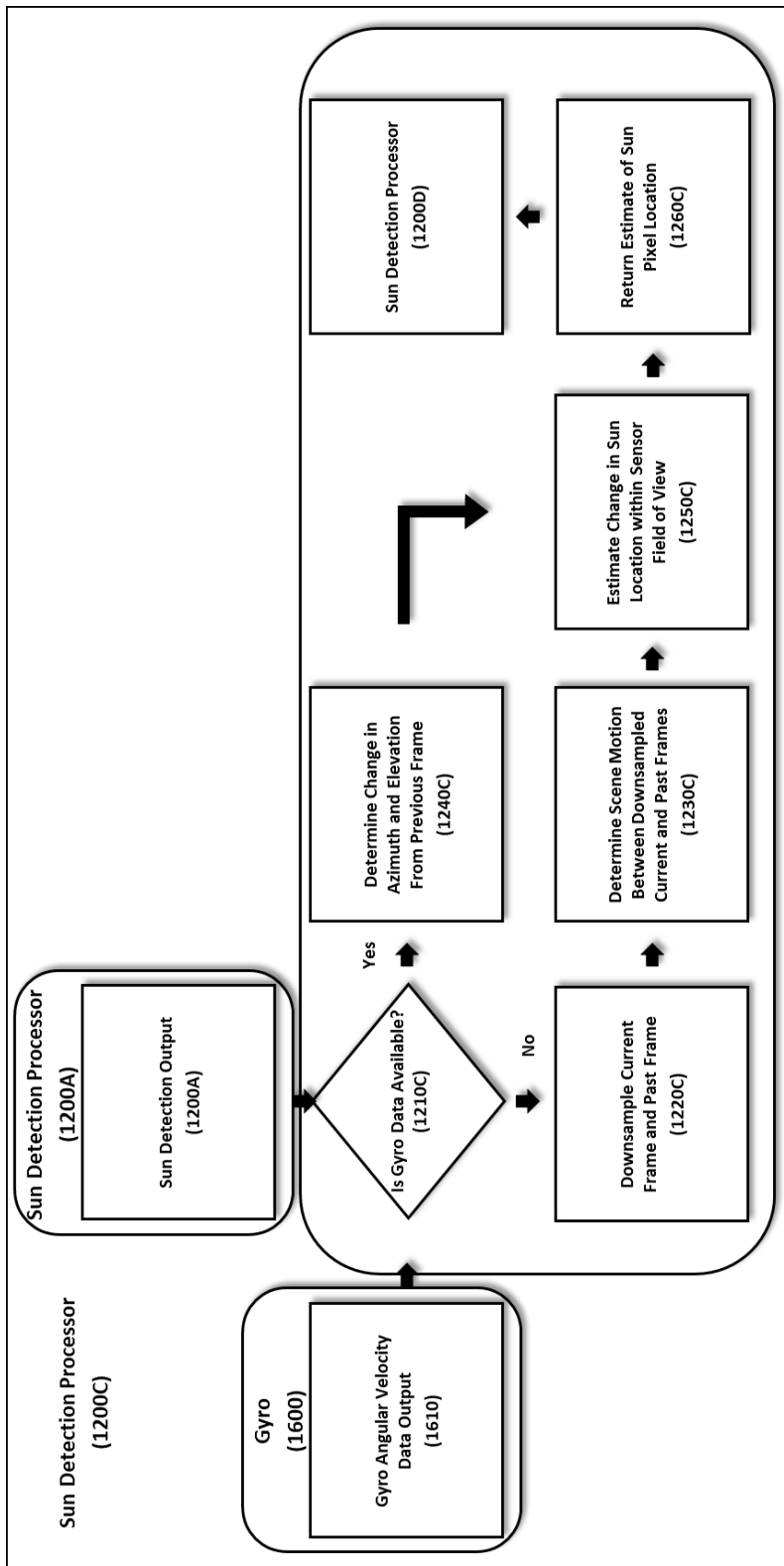


Figure IV.8. Sun Tracking

Discussion

The designs and associated methods presented in chapter provide techniques for mitigating saturation and clutter which are not fully addressed in chapters 2 and 3. The designs presented in this chapter attempt to eliminate saturation and clutter by use of additional hardware and accompanying algorithmic processes. As a baseline, an additional detection sensor band is added to reduce false alarms. Three additional options are provided to lessen the extent of saturation clutter. A dedicated imaging sensor for saturation events adds a greater amount of SWAP-C, but also provides increased performance without latency. An acoustic saturation sensor option delivers increased performance with lower increase in SWAP-C, at the expense of latency added to the declaration time. The fourth option using orientation, position, and time offers increased performance, with minimal impact to SWAP-C, but may only be applicable in environments where this data is available. Design options are provided for varying levels of performance associated with clutter mitigation, saturation mitigation, SWAP-C, processing capability, and available data.

V. CONCLUSIONS, CONTRIBUTIONS, AND FUTURE WORK

Conclusion

This work presents novel methods for implementing a practical impulse detection system, which has been applied to weapon fire detection. The methodology incorporates multiple layers of techniques to determine detection and provides a custom classification scheme, which improves classification of target classes and false alarms. Real time processing and data management techniques are included in the methodology. The work also emphasizes the pairing between an algorithm and system in which it resides. Multiple systems designs are presented which address limitations in the algorithms. Similar to algorithm techniques, a single solution does not address every issue, and pros and cons of each system are provided, along with suggestions for additional capability.

Contributions

The algorithm and system designs take into account real-world implementation requirements. These requirements include detection performance, target classification, false alarm and clutter mitigation, cost, range performance, hardware availability, real-time processing, on-the-move operation, and multi-use applications. The work in this dissertation established and verified weapon fire detection at low frame rates, and practical implementation in various wavebands, due to clutter mitigation techniques. The novelty of the proposed dissertation has been recognized by the United States Patent and Trademark Office, by the acceptance and publishing of two patents resulting from this work. The patents emphasized

adaptive clutter mitigation strategies, and also emphasized saturation clutter mitigation strategies and systems.

Future Work

Potential future work can include phenomenology and signature analysis, incorporation of novel sensors, implementation across multiple processing architectures, and full demonstration of on-the move applications. Phenomenology and signature analysis can exploit the intensity of radiated energy in various wavebands, taking into account, atmospheric attenuation, to improve signal to noise, affecting detection and range performance. Novel sensors may improve upon available sensor formats and performance, for increased number of pixels and increased sensitivity, at varying frame rates. The availability of these sensors in different wavebands, can ultimately affect system design. Additionally, the cost of new, and even older sensors can also affect system design. For example, inexpensive sensors may allow the problem space to be decomposed, such that each sensor(s) is addressing a different aspect of the problem, possibly reducing algorithm complexity and sensor performance requirements. Implementation, using various processing architectures, should be investigated to determine the trade space of performance, real-time processing, and reduced size and power. This research is critical for enabling low cost, efficient mobile and remote processing, and could increase commercial applicability. Additionally, demonstration of acceptable on-the-move performance could also increase the number of commercial applications. Commercial applications could include gun fire and explosive detonation detection. Commercial, automated remote sensing can be used in high crime areas, schools, or at large public events. Additionally, police vehicles can be equipped with sensors enabling autonomous gun fire detection, imaging, and reporting. Further research is needed to investigate modifications to the algorithm for mobile application, investigate

incorporation of image stabilization and registration techniques and possible associated hardware, and how these modifications affect detection performance, processing, size, power and cost.

REFERENCES

- [1] N. A. MacMillan and D. C. Creelman, *Detection Theory" A User's Guide*, Mahwah: Lawrence Erlbaum Associates, 2005.
- [2] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, New York: John Wiley & Sons, Inc, 2001.
- [3] S. Zuffi, C. Brabilla, G. Beretta and P. Scala, "Human Computer Interaction: Legibility and Contrast," *Image Analysis and Processing*, vol. 14th International Conference on Image Analysis and Processing, pp. 241-246, 2007.
- [4] E. Peli, "Contrast in complex images," *J. Opt. Soc. Am.*, vol. 7, no. 10, p. 2032, 1990.
- [5] M. Stone, "Contrast Metrics Explained," [Online]. Available: <http://www.stonesc.com/pubs/Contrast%20Metrics.htm>. [Accessed 8 10 2013].
- [6] J. L. Rowe, "The Impact of Thermal Imaging Camera Display Quality on Fire Fighter Task Performance," The University of Maryland, College Park, 2008.
- [7] V. A. Hodgkin, T. Maurer, C. Halford and R. Vollmerhausen, "Impact of Path Radiance on MWIR and LWIR Imaging," in *Proc. of SPIE. Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XVIII*, Orlando, 2007.
- [8] F. A. Smith, E. L. Jacobs, S. Chari and J. Brooks, "LWIR thermal imaging through dust obscuration," in *Proc. of SPIE. Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXII*, Orlando, 2011.
- [9] C. A. Schuetz, E. L. Stein Jr., J. Samluk, D. MacKrides, J. P. Wilson, R. D. Martin, T. E. Dillon and D. W. Prather, "Studies of Millimeter-wave Phenomenology for Helicopter Brownout Mitigation," in *Proc. of SPIE. Millimetre Wave and Terahertz Sensors and Technology II*, Berlin, 2009.
- [10] J. B. Brown, S. Chari, J. Hutchison, J. Gabonia and E. Jacobs, "Assessment of a linear pyroelectric array sensor for profile classification," in *Unattended Ground, Sea, and Air Sensor Technologies and Applications XII, Proc of SPIE*, Orlando, 2010.
- [11] J. B. Brown, "Thesis: Signal Conditioning and Feature Estimation for Profiling Sensor Systems," The University of Memphis, Memphis, 2010.

- [12] D. D. Baumgartner and B. J. Schachter, "Improving FLIR ATR Performance in a Turbulent Atmosphere with a Moving Platform," in *Automatic Target Recognition XXII, Proc of SPIE*, Orlando, 2012.
- [13] D. D. Baumgartner, J. B. Brown, E. L. Jacobs and B. J. Schachter, "Degraded visual environment image / video quality metrics," in *Degraded Visual Environments: Enhanced, Synthetic, and External Vision Solutions, Proc of SPIE*, Baltimore, 2014.
- [14] J. B. Brown, S. K. Chari and E. L. Jacobs, "An assessment of a 360 degree profiling sensor for object classification," in *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR II, Proc of SPIE*, Orlando, 2011.
- [15] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977-1000, 2003.
- [16] S. S. Streetman and M. W. McGarvey, "Method and Device for Event Detection Utilizing Data From A Multiplicity of Sensor Sources". United States of America Patent 2002/0196140, 26 December 2002.
- [17] G. Klingenberg and J. M. Heimerl, Gun Muzzle Blast and Flash, vol. 139, A. R. Seebass, Ed., Washington DC: American Institute of Aeronautics and Astronautics, 1992, p. 97.
- [18] R. G. Driggers, P. Cox and T. Edwards, Introduction to Infrared and Electro-Optical Systems, Boston: Artech House, Inc, 1999.
- [19] F. S. Simmons, Rocket Exhaust Plume Phenomenology, El Segundo: The Aerospace Press, 2000, p. 157.
- [20] G. Tidhar, "Gunshot Detection System and Method". United States of America Patent 2011/0170798, 22 July 2010.
- [21] G. Tihar, "Gunshot Detection System and Method". United States of America Patent 2014/0184806, 16 September 2013.
- [22] G. Tidhar, "Multiband Flash Detection System". United States of America Patent 2015/0138534, 5 November 2014.
- [23] R. C. Maher, "Acoustical Characterization of Gunshots," in *IEEE Workshop on Signal Processing*, Washington D.C., 2007.
- [24] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," in *IEEE 12th Digital Signal Processing Workshop*, Jackson Lake, 2006.

- [25] R. C. Maher and S. R. Shaw, "Deciphering Gunshot Recordings," in *AES 33rd International Conference*, Denver, 2008.
- [26] M. L. Johnson, "Systems and Methods of Processing Information Regarding Weapon Fire location Using Projectile Shockwave and Muzzle Blast Times of Arrival Data". United States of America Patent 8,995,227, 15 August 2011.
- [27] I. A. Mantilla-Gaviria, M. Leonardi, G. Galati and J. V. Balbastre-Tejedor, "Localization algorithms for multilateration (MLAT) systems in airport surface surveillance," *Signal, Image and Video Processing*, vol. 9, no. 7, pp. 1549-1558, 2015.
- [28] J. R. Aguilar, "Gunshot Location Systems: The transfer of the sniper detection technology from military to civilian applications," in *Conference: 47th International Carnahan Conference on Security Technology (ICCST)*, Medellin, Colombia, 2013.
- [29] A. Chacon-Rodriguez, P. Julian, L. Castro, P. Alvarado and N. Hernandez, "Evaluation of Gunshot Detection Algorithms," *IEEE Transactions on Circuits and Systems*, vol. 58, no. 2, pp. 363-373, 2011.
- [30] Department of Defense, Department of Homeland Security, Department of Transportation, "2017 Federal Radionavigation Plan," Department of Defense, Department of Homeland Security, Department of Transportation, Springfield, VA, 2017.
- [31] National Coordination Office for Space-Based Positioning, Navigation, and Timing, "GPS.GOV," National Oceanic and Atmospheric Administration, [Online]. Available: <https://www.gps.gov/systems/gps/performance/accuracy/>. [Accessed 23 May 2020].
- [32] I. Obeid and P. D. Wolf, "Evaluation of Spike-Detection Algorithms for a Brain-Machine Interface Application," *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, vol. 51, no. 6, pp. 905-911, June 2004.
- [33] A. Zviaginstev, Y. Perelman and R. Ginosar, "Algorithms and Architectures for Low Power Spike Detection and Alignment," *Journal of Neural Engineering*, vol. 3, no. 1, pp. 35-42, 20 January 2006.
- [34] Trimble, "GPS Tutorial," Trimble, [Online]. Available: https://www.trimble.com/gps_tutorial/. [Accessed 23 May 2020].
- [35] MapToaster, "MapToaster Topo: How GPS Works," Integrated Mapping Ltd, 2014. [Online]. Available: <https://www.maptoaster.com/maptoaster-topo-nz/articles/how-gps-works/how-gps-works.html>. [Accessed 24 May 2020].

- [36] J. B. Brown, J. E. Hutchison III, J. K. Gabonia, J. H. Davis, "Weapon Fire Detection and Localization System for Electro-Optical Sensors". U.S. Patent US 10,209,343 B1, 2019
- [37] J. B. Brown, J. E. Hutchison III, "Weapon Fire Detection and Localization Algorithm for Electro-Optical Sensors". U.S. Patent US 10,389,928 B2, 2019.
- [38] J. B. Brown, J. E. Hutchison III, C. L. Howell "Weapon Fire Detection, Emphasizing Clutter and Saturation Mitigation." Journal of DoD Research & Engineering, TBD(TBD), 2022.
- [39] J. B. Brown, J. E. Hutchison III, C. L. Howell, J. K. Gabonia, E. L. Jacobs, "Advancement in Weapon Fire Detection Methodologies." Journal of DoD Research & Engineering, TBD(TBD), 2022.